

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:47:40 ; Search time 32.57 Seconds
(without alignments)
19,831 Million cell updates/sec

Title: US-08-981-824-1

Perfect score: 110
Sequence: 1 DVNATFLHATDLPACDGER 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 1436

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	23.6	19	1	PHSL_DESBN
2	25	22.7	23	1	PAP2_MANSE
3	24	21.8	12	1	XYLA_STRVN
4	24	21.8	18	1	HSTB_ECOLI
5	24	21.8	23	1	PAP1_MANSE
6	23	20.9	10	1	FARP_LOEMI
7	23	20.9	10	1	LCMS_LEUMA
8	23	20.9	10	1	NEMS_DROME
9	23	20.9	19	1	DHAB_COMTE
10	22	20.0	19	1	DURC_STRGP
11	22	20.0	19	1	LANC_STRSQ
12	22	20.0	24	1	CH60_HELVI
13	22	20.0	25	1	RS11_TRETH
14	21	19.1	19	1	DURA_STRGV
15	21	19.1	19	1	DURE_STRGV
16	21	19.1	22	1	CH60_BOVIN
17	20.5	18.6	13	1	ACT7_SOYBN
18	20.5	18.6	17	1	ACT6_SOYBN
19	20.5	18.6	25	1	FLA4_TREPH
20	20	18.2	10	1	FARP_MANSE
21	20	18.2	13	1	FXA1_CONGE
22	20	18.2	13	1	FXA2_CONGE
23	20	18.2	13	1	FXA1_CONST
24	20	18.2	14	1	FXA1_CONCN
25	20	18.2	14	1	FXA1_CONMA
26	20	18.2	15	1	SODE_PINPS
27	20	18.2	16	1	HTPC_ACICA
28	20	18.2	19	1	FXA2_CONST
29	20	18.2	20	1	STVA_STYCL
30	20	18.2	20	1	STVB_STYCL
31	20	18.2	20	1	THIO_CANFA
32	20	18.2	20	1	FIBB_CEREL
33	20	18.2	21	1	FIBB_CEREL

34	20	18.2	21	1	FIBB_MUNMU
35	20	18.2	21	1	NRLA_ACISP
36	20	18.2	25	1	BOTR_BOTJA
37	20	18.2	25	1	CYCA_VIBFI
38	19.5	17.7	22	1	HG42_FASHE
39	19	17.3	13	1	UP51_UPEIN
40	19	17.3	14	1	MARI_ALTSP
41	19	17.3	14	1	SAP2_ARBPV
42	19	17.3	15	1	UC30_MAIZE
43	19	17.3	16	1	BA11_EUBSP
44	19	17.3	16	1	FXA1_CONPE
45	19	17.3	16	1	FXA1_CONPE

ALIGNMENTS

RESULT	1	PHSL_DESBN	STANDARD	PRT	19 AA
AC	P13066				
DT	01-JAN-1990 (Rel. 13, Created)				
DT	01-JAN-1990 (Rel. 13, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	PERIPLASMIC [NIFESE] HYDROGENASE LARGE SUBUNIT (EC 1.18.99.1) (NIFESE DE HYDROGENYLASE LARGE CHAIN) (FRAGMENT).				
OS	Desulfovibrio baculatus (strain Norway 4).				
OC	Bacteria; Proteobacteria; delta subdivision; Desulfomicrobium.				
RN	[1]				
RP	SEQUENCE.				
RX	MEDLINE-88106446; PubMed-3322275;				
RA	Pickrell B.C., He S.H., Li C., Menon N., Choi E.S., Przybyla A.E.,				
RA	Devartanian D.V., Peck H.D. Jr., Faugue G., le Gall J., Teixeira M.,				
RA	Moura I., Moura J.J.G., Patil D., Huynh B.H.;				
RT	Identification of three classes of hydrogenase in the genus,				
RT	Desulfovibrio.;				
RT	Biochem. Biophys. Res. Commun. 149:369-377(1987).				
CC	-1- CATALYTIC ACTIVITY: 2 REDUCED FERREDOXIN + 2 H(+) -> 2 OXIDIZED				
CC	-1- FERREDOXIN + H(2).				
CC	-1- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CLUSTERS AND ONE				
CC	SELENOCYSTEINE.				
CC	-1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.				
CC	-1- SUBCELLULAR LOCATION: PERIPLASMIC.				
CC	-1- SIMILARITY: BELONGS TO THE [NIFE]/[NIFESE] HYDROGENASE LARGE				
DR	PIR; H27480; H27480.				
DR	INTERPRO: IPR001501; .				
DR	PROSITE: PS00507; NI_HGNASE_L.1; PARTIAL.				
KW	Oxidoreductase; Periplasmic; Nickel; Selenium; Selenocysteine.				
FT	NON_TER				
SQ	SEQUENCE 19 AA; 1942 MW; 2BFCDD2D360F00367 CRC64;				

Query Match Score 26; DB 1; Length 19;

Best Local Similarity 57.1%; Pred. No. 3.4e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 14 PACDGER 20

DB 6 PADGKK 12

RESULT 2

PAP2_MANSE STANDARD; PRT; 23 AA.

AC P30254; .

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE PARALYTIC PEPTIDE II (PP II).

OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

CC	Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC	Sphingioidea; Sphingidae; Sphinginae; Manduca.
RP	[1]
RN	SEQUENCE.
BP	TISSUE=HEMOLYMPH.
RC	MEDLINE=91302298; PubMed-2071576;
RX	Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA	Quistad G.B.;
RA	"Isolation and identification of paralytic peptides from hemolymph of
RT	the lepidopteran insects <i>Manduca sexta</i> , <i>Spodoptera exigua</i> , and
RT	<i>Heliothis virescens</i> ."
RT	J. Biol. Chem. 266:12873-12877(1991).
RL	- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC	LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC	HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC	- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR	PIR: B39855; B39855.
KW	Hemolymph.
FT	DISUFD 7 19 BY SIMILARITY.
SO	SEQUENCE 23 AA; 2420 MW; 0B26GB5C31263FE4 CRC64;

Query Match	22.7%	Score 25;	DB 1;	Length 23;
Best Local Similarity	43.8%	Pred. No. 5;	get02;	
Matches	7;	Conservative	1;	Mismatches 8;
				Indels 0;
				Gaps 0;
QY	3	NTAFLHATDLPACDG	18	
	1:1	11111111		
Db	2	NFAGGCATGFLRTADG	17	

	RESULT	3			
ID	XYLA_STRVN		STANDARD;	PRT;	12 AA.
AC	P14405:				
DT	01-JAN-1990 (Rel. 13, Created)				
DT	01-JAN-1990 (Rel. 13, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	XYLOSE ISOMERASE [EC 5.3.1.5] (FRAGMENT).				
OS	Streptomyces violaceoruber..				
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;				
CC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.				
RN	(1)				
RP	SEQUENCE.				
RC	STRAIN-LMG 7183;				
RX	MEDLINE=90104230; PubMed=2604694;				
RA	Vangiersperre W., Ampe C., Kersters-Hilderson H., Tempst P.;				
RT	"Single active-site histidine in D-xylase isomerase from Streptomyces				
RT	violaceoruber. Identification by chemical derivatization and peptide				
RT	mapping.";				
RL	Biochem. J. 263:195-199(1989).				
CC	- - CATALYTIC ACTIVITY: D-XULOSE -> D-XYLUOSE.				
CC	- - COPACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.				
CC	- - SUBUNIT: HOMOTETRAMER.				
CC	- - SUBCELLULAR LOCATION: CYTOPLASMIC.				
CC	- - SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.				
HSSP	P37031: IDXI.				
DR	INTERPRO: IPR001998: "				
DR	PROSITE: PS00172: XYLOSE_ISOMERASE_1; PARTIAL.				
DR	PROSITE: PS00173: XYLOSE_ISOMERASE_2; PARTIAL.				
KW	Isomerase; Pentose shunt; Xylose metabolism; Magnesium.				
FT	NON_TER	1			
FT	ACT_SITE	5			
FT	NON_TER	12			
FT	NON_TER	12			
EQ	SEQUENCE	12 AA; 1375 MW; E749266EB1AAAAA1 CRC64;			

Query Match	21.8%	Score 24;	DB 1;	Length 12;
Best Local Similarity	57.1%	Pred. No. 4.5e+02;		
Matches 4;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0;
QY	8	HATDLP 14		
		:		

```

DB      5  HDDLIP 11

RESULT      4
HSTB_ECOLI  STANDARD;  PRT;  18 AA.
AC      P01560;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      15-DEC-1998 (Rel. 37, Last annotation update)
DE      HEAT-STABLE ENTEROTOXIN ST-2 (ST-B).
OC      Escherichia coli.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
RN      [1]
RP      SEQUENCE.
RC      STRAIN-18D / SEROTYPE 0.43;K66:H37;
RX      MEDLINE-81264141; PubMed=7021541;
RA      Chan S.-K., Giannelis R.A.;
RT      "Amino acid sequence of heat-stable enterotoxin produced by
RL      Escherichia coli pathogenic for man.";
RN      J. Biol. Chem. 256:7744-7746(1981).
RN      [2]
RP      DISULFIDE BONDS.
RX      MEDLINE-87191003; PubMed=3552731;
RA      Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Aimoto S., Takeda T.,
RT      Miwatani T., Takeda Y.;
RT      "Mode of disulfide bond formation of a heat-stable enterotoxin (StH)
RT      produced by a human strain of enterotoxigenic Escherichia coli.";
RN      FEBS Lett. 215:165-170(1987).
CC      -1- FUNCTION: TOXIN WHICH ACTIVATES THE PARTICULATE FORM OF GUANYLATE
CC      CYCLASE AND INCREASES CYCLIC GMP LEVELS WITHIN THE HOST
CC      INTESTINAL EPITHELIAL CELLS.
CC      -1- DISEASE: BOTH HEAT-STABLE AND HEAT-LABILE ENTEROTOXINS ARE
CC      PRODUCED BY PARATHOGENIC STRAINS OF E.COLI AND AFFECT THE DIGESTIVE
CC      TRACT OF MAMMALS.
CC      -1- SIMILARITY: BELONGS TO THE HEAT-STABLE ENTEROTOXIN FAMILY.
DR      PIR: A01893; QHNC2.
DR      HSSP; P01559; IETN.
DR      INTERPRO; IPR001489; -.
DR      PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
KW      Toxin; Enterotoxin.
FT      DISULFID 5 10
FT      DISULFID 6 14
FT      DISULFID 9 17
FT      DISULFID 18 AA; 1978 MW; DDC975F49D600650 CRC64;
SQ      SEQUENCE

Query Match      21.8%; Score 24; DB 1; Length 18;
Best Local Similarity 80.08; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      14  PACDQ 18
DB      12  PACAG 16

RESULT      5
PAP1_MANSE  STANDARD;  PRT;  23 AA.
AC      P30253;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      PARALYTIC PEPTIDE I (Pe I).
DE      Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC      Euteleyota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC      Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Dictyssa;
OC      Sphingioidea; Sphingidae; Sphinginae; Manduca.
RN      [1]
RP      SEQUENCE.
RC      TISSUE=HEMOLYMPH;
RX      MEDLINE-91302298; PubMed=2071576;
```

RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
 RA Quistad G.B.:
 RT "Isolation and identification of paralytic peptides from hemolymph of
 RT the lepidopteran insects *Manduca sexta*, *Spodoptera exigua*, and
 RT *Heliothis virescens*.";
 RL J. Biol. Chem. 266:12873-12877(1991).
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
 CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
 CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 DR PIR: A39855; A39855.
 CC Hemolymph.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA: 2436 MW: 0B26CB5C29855FE4 CRC64;

Query Match 21.8%; Score 24; DB 1; Length 23;
 Best Local Similarity 43.8%; Pred. No. 8.6e+02;
 Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 3 NYAFHATDLPACDG 18
 Db 2 NFAGCATGYLRTADG 17

RESULT 6
 FARP.LOCMI STANDARD; PRT; 10 AA.
 AC P38553;
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SCHISTOFLR-AMIDE (PDVDFLRF-AMIDE) (CARDIOEXCITATORY NEUROPEPTIDE).
 OS Locustocerca gregaria (Migratory locust), and
 OS Schistocerca gregaria (Desert locust).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
 OC Acridomorpha; Acridoidea; Acrididae; Locusta.
 RN [1]
 RP SEQUENCE.
 RC SPECIES-S.GREGARIA; TISSUE-THORACIC NERVOUS SYSTEM;
 RX MEDLINE-93324430; PubMed-7687352;
 RA Schoofs L., Holman G.M., Paemen L., Veeleert D., Amelknock M.,
 RA de Loof A.:
 RT "Isolation, identification, and synthesis of PDVDFLRFamide
 RT (Schistoflrfamide) in *Locusta migratoria* and its association with the
 RT male accessory glands, the salivary glands, the heart, and the
 RT oviduct.";
 RL Peptides 14:409-421(1993).
 RN [2]
 RP SEQUENCE.
 RC SPECIES-S.GREGARIA; TISSUE-THORACIC NERVOUS SYSTEM;
 RX MEDLINE-89246543; PubMed-2719702;
 RA Robb S., Packman L.C., Evans P.D.:
 RT "Isolation, primary structure and bioactivity of schistoflrfamide, a
 RT FMRF-amide-like neuropeptide from the locust, *Schistocerca*
 RT gregaria.";
 RL Biochem. Biophys. Res. Commun. 160:850-856(1989).
 CC -1- FUNCTION: MUSCLE INHIBITING AGENT. INVOLVED IN THE NEURAL CONTROL
 CC OF THE VISCERAL MUSCLES OF THE HEART, ACCESSORY GLANDS AND
 CC OVIDUCT. MAY BE INVOLVED IN THE REGULATION OF SALIVA SECRETION.
 CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE MALE ACCESSORY GLANDS,
 CC THE SALIVARY GLANDS, THE HEART, AND THE OVIDUCT.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR PIR: A32543; A32543.
 CC Neuropeptide: Amidation.
 FT MOD.RES 10 10
 SQ SEQUENCE 10 AA: 1244 MW: D3C51729D2C1EAB2 CRC64;

Query Match 20.9%; Score 23; DB 1; Length 10;
 Best Local Similarity 57.1%; Pred. No. 5.5e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 DVNYAFL 7
 Db 2 DVDHVF 8

RESULT 7
 LCMS.LEUMA STANDARD; PRT; 10 AA.
 AC P21144; P41497;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE LEUCOMYOSUPPRESSIN (LMS) (LEM-MS).
 OS Leucophaea maderae (Madeira cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberidae; Blaberidae; Leucophaea.
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE-HEAD;
 RA Holman G.M., Cook B.J., Nachman R.J.:
 RT "Isolation, primary structure and synthesis of leucomyosuppressin,
 RT an insect neuropeptide that inhibits spontaneous contractions of the
 RT cockroach hindgut.";
 RL Comp. Biochem. Physiol. 85C:329-333(1986).
 CC -1- FUNCTION: INHIBITS THE SPONTANEOUS CONTRACTIONS OF COCKROACH
 CC PROTEIN (HINDGUT).
 CC Neuropeptide; Amidation.
 FT MOD.RES 1 1
 FT MOD.RES 10 10
 SQ SEQUENCE 10 AA: 1275 MW: D3C45229D2C1EAB2 CRC64;

Query Match 20.9%; Score 23; DB 1; Length 10;
 Best Local Similarity 57.1%; Pred. No. 5.5e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DVNYAFL 7
 Db 2 DVDHVF 8

RESULT 8
 NEMS.DROME STANDARD; PRT; 10 AA.
 AC P41494;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE NEOMYOSUPPRESSIN (NEB-MS).
 GN NEMS.
 OS Drosophila melanogaster (Fruit fly), and
 OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE.
 RC SPECIES-D.MELANOGASTER;
 RX MEDLINE-93002195; PubMed-1390001;
 RA Nichols R.:
 RT "Isolation and structural characterization of Drosophila
 RT TDVDFLRFamide and FMRFamide-containing neural peptides.";
 RL J. Mol. Neurosci. 3:213-218(1992).
 RN [2]
 RP SEQUENCE.
 RC SPECIES-S.BULLATA; TISSUE-HEAD;
 RX MEDLINE-93047886; PubMed-1358537;
 RA Fonagy A., Schoofs L., Proost P., Van Damme J., Bueds H., De Loof A.:
 RT "Isolation, primary structure and synthesis of neomyosuppressin, a
 RT myoinhibiting neuropeptide from the grey fleshfly, *Neobellieria*

```

RT bullata."
RL Comp. Biochem. Physiol. 102C:239-245(1992).
CC -1- FUNCTION: MYOINHIBITING NEUROPEPTIDE.
DR FLVASE; FBgn0013996; Nems.
KM Neuropeptide; Amidation.
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1248 MW; D3C00329D2C1EAB2 CRC64;

Query Match
Best Local Similarity 20.9%; Score 23; DB 1; Length 10;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DVNTAF 7
Db 2 DVNTAF 8

RESULT 9
DHAB_COMTE STANDARD; PRT; 19 AA.
ID DHAB_COMTE
AC P80704;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ALDEHYDE DEHYDROGENASE, BETA CHAIN (EC 1.2.99.3) (ALDH) (FRAGMENT).
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
RN [1]
RP SEQUENCE.
RC STRAIN-ATCC 15667;
RA Luykx D.M.A.M., Kim S.W., de Vries S., Duine J.A.;
RL Submitted (JUL-1996) to the SWISS-PROT data bank.
CC -1- CATALYTIC ACTIVITY: AN ALDEHYDE + ACCEPTOR + H(2)O = AN ACID +
CC REDUCED ACCEPTOR.
CC -1- COFACTOR: MOLYBDENUM.
CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA AND A GAMMA
CC CHAIN.
KM Oxidoreductase; Molybdenum.
FT NON_TER 19
SQ SEQUENCE 19 AA; 2096 MW; 89BD67D4D05A212E CRC64;

Query Match
Best Local Similarity 20.9%; Score 23; DB 1; Length 19;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 YAFUAT 10
Db 2 YAFYST 8

RESULT 10
DURC_STRGP STANDARD; PRT; 19 AA.
ID DURC_STRGP
AC P36503;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE LANTIBIOTIC DURAMYCIN C.
OS Streptomyces griseolutes.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE.
RC STRAIN-R2107;
RA MEDLINE-91107436; PubMed-2125590;
RA Friedenham A., Fendrich G., Marki F., Gruner J.,
RA Raschdorf F., Peter H.H., Marki W., Gruner J.,
RT "Duramycin B and C, two new lanthionine containing antibiotics as
RT inhibitors of phospholipase A2. Structural revision of duramycin and
RT cinnamycin."
RL J. Antibiot. 43:1403-1412(1990).

```

```

RN [2]
RP STRUCTURE BY NMR.
RA Zimmerman N., Freund S., Friedenham A., Jung G.;
RT "Solution structure of the lantibiotics duramycin B and C.";
RL (in) Schneider C.H., Eberle A.N. (eds.);
RL Peptides 1992, pp.519-520, Escom Science Publishers, Leiden (1992).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE-93387292; PubMed-8375380;
RA Zimmerman N., Freund S., Friedenham A., Jung G.;
RT "Solution structures of the lantibiotics duramycin B and C.";
RL Eur. J. Biochem. 216:419-428(1993).
CC -1- FUNCTION: ACTS AS INHIBITOR OF PHOSPHOLIPASE A2.
CC -1- PTM: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF
CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE
CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND
CC CLEAVAGE OF THE MODIFIED PRECURSOR.
CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE B LANTIBIOTICS.
KM Antibiotic; Lantibiotic.
FT MOD_RES 4
FT MOD_RES 6
FT MOD_RES 11
FT MOD_RES 11
FT MOD_RES 18
FT MOD_RES 18
FT MOD_RES 19
FT MOD_RES 19
FT THIOETH 4
FT THIOETH 4
FT THIOETH 11
SQ SEQUENCE 19 AA; 2007 MW; E2404E3E3F5286A CRC64;

Query Match
Best Local Similarity 20.0%; Score 22; DB 1; Length 19;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 ACDE 18
Db 13 SCDE 16

RESULT 11
LANC_STRSQ STANDARD; PRT; 19 AA.
ID LANC_STRSQ
AC P38655;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE LANTIBIOTIC ANCOVENIN.
OS Streptomyces sp. (strain A647P-2).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE.
RA Wakamaya T., Ueki Y., Shiba T., Kido Y., Motoki Y.;
RT "The structure of ancovenin, a new peptide inhibitor of angiotensin I
RT converting enzyme.";
RL retrieved from Lett. 26:665-668(1985).
CC -1- FUNCTION: ACTS AS INHIBITOR OF ANGIOTENSIN I CONVERTING ENZYME.
CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE B LANTIBIOTICS.
DR PIR; A61284; EWSMAN.
KM Antibiotic; Lantibiotic.
FT MOD_RES 4
FT MOD_RES 6
FT MOD_RES 11
FT MOD_RES 11
FT MOD_RES 18
FT MOD_RES 18
FT MOD_RES 19
FT MOD_RES 19
FT THIOETH 4
FT THIOETH 4
FT THIOETH 11
SQ SEQUENCE 19 AA; 2033 MW; F434299E2736286A CRC64;

Query Match
Best Local Similarity 20.0%; Score 22; DB 1; Length 19;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```


Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 15 ACDB 18
:111
Db 13 SCDB 16

RESULT 12

CH60_HELVI
ID CH60_HELVI STANDARD; PRT; 24 AA.
AC P26317;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 60 KDA CHAPERONIN, MITOCHONDRIAL (P60) (FRAGMENT).
OS Heliothis virescens (Noctuid moth) (Owlet moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Plekoyota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Heliothinae; Heliothis.

SEQUENCE.
RC TISSUE-TESTIS;
RC MEDLINE-90339485; PubMed-1974308;
RA Miller S.G., Leclerc R.F., Erdos G.W.;
RT "Identification and characterization of a testis-specific isoform of a chaperonin in a moth, Heliothis virescens.";
RL J. Mol. Biol. 214:407-422(1990).
CC -1- FUNCTION: IMPLICATED IN MITOCHONDRIAL PROTEIN IMPORT AND MACROMOLECULAR ASSEMBLY. MAY FACILITATE THE CORRECT FOLDING OF IMPORTED PROTEINS. MAY ALSO PREVENT MISFOLDING AND PROMOTE THE REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS CONDITIONS IN THE MITOCHONDRIAL MATRIX (BY SIMILARITY).
CC -1- SUBUNIT: FORMS A SINGLE SEVEN-MEMBER RING COMPLEX, IN TIGHT ASSOCIATION WITH THE P63 PROTEIN.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DEVELOPMENTAL STAGE: FROM THE SECOND HALF OF THE LARVAL FINAL-INSTAR, THROUGH THE FIRST TWO DAYS OF PUPAL DEVELOPMENT.
CC -1- MISCELLANEOUS: SHOWS ATPASE ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
DR HSSP; P06139; IAOB.
DR INTERPRO: IPR001844;
DR PROSITE: PS00296; CHAPERONINS_CPN60; PARTIAL.
KM Chaperone: ATP-binding; Testis; Mitochondrion.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2531 MW; 2B34508F8CA981CF CRC64;

Query Match 20.0%; Score 22; DB 1; Length 24;
Best Local Similarity 38.5%; Pred. No. 1.9e+03;
Matches 5; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 DVNYAFLLATDIL 13
:111
Db 9 DVRLMLQGVVL 21

RESULT 13

RS11_THERH
ID RS11_THERH STANDARD; PRT; 25 AA.
AC P80376;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S11 (FRAGMENT).
GN RPSK OR RPS11.
OS Thermus aquaticus (subsp. thermophilus).
OC Bacteria; Thermus/Delnochococcus group; Thermus group; Thermus.

SEQUENCE.
RX MEDLINE-95045586; PubMed-7957245;
RA Tsiiboli P., Herfurth E., Choi T.;

RT "Purification and characterization of the 30S ribosomal proteins from the bacterium Thermus thermophilus.";

RT Eur.J. Biochem. 226:169-177(1994).
CC -1- FUNCTION: S11 PLAYS AN ESSENTIAL ROLE FOR THE SELECTION OF THE CORRECT TRNA IN PROTEIN BIOSYNTHESIS. IT IS LOCATED ON THE LARGE LOBE OF THE SMALL SUBUNIT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S11P FAMILY OF RIBOSOMAL PROTEINS.
DR INTERPRO: IPR001971;
DR PROSITE: PS00054; RIBOSOMAL_S11; PARTIAL.
KM Ribosomal protein.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2816 MW; 59A3F932857D88F CRC64;

Query Match 20.0%; Score 22; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 AFLHAT 10
:1111
Db 18 AYTHAS 23

RESULT 14

DURA_STRGV
ID DURA_STRGV STANDARD; PRT; 19 AA.
AC P36504;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE LANTIBIOTIC DURAMYCIN (LEUCOPEPTIN) (ANTIBIOTIC PA48009).
OS Streptococcus salivarius griseovercillium.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae;
OC Streptovercillium.
RN [1]
RP SEQUENCE AND STRUCTURE BY NMR.
RC STRAIN-PA-48009;
RX MEDLINE-91107438; PubMed-2272918;
RA Hayashi F., Nagashima K., Terui Y., Kawamura Y., Matsumoto K., Itazaki H.;
RT "The structure of PA48009: the revised structure of duramycin.";
RL J. Antibiot. 43:1421-1430(1990).
RN [2]

RP SEQUENCE.
RX MEDLINE-91107436; PubMed-2125590;
RA Fiedenhagen A., Fendrich G., Markl F., Markl W., Gruner J., Raschdorf F., Peter H.H.;
RT "Duramycin B and C, two new lantibiotics containing antibiotics as inhibitors of phospholipase A2. Structural revision of duramycin and lantibiotin.";
RL J. Antibiot. 43:1403-1412(1990).
CC -1- FUNCTION: ACTS AS INHIBITOR OF PHOSPHOLIPASE A2.
CC -1- PTM: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF THR, SER, AND CYS INTO DEHYDROXY AA AND THE FORMATION OF SULFIDE BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND CLEAVAGE OF THE MODIFIED PRECURSOR.
CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE B LANTIBIOTICS.
KM Antidiotic; Lantibiotic.
FT MOD_RES 4 4
FT MOD_RES 6 6
FT MOD_RES 11 11
FT MOD_RES 18 18
FT MOD_RES 19 19
FT THIOETH 1 18
FT THIOETH 4 14
FT THIOETH 5 11
SQ SEQUENCE 19 AA; 2069 MW; 012951AE27362F00 CRC64;

Query Match 19.1%; Score 21; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 CDG 18
 DB 14 CDG 16

RESULT 15

DURB_STRGM STANDARD: PRT: 19 AA.

AC P36502;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE LANTHBIOTIC DURAMYCIN B.
 OS Streptovorticillium sp. (strain R2075).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae;
 OC Streptovorticillium.
 CC [1]
 RP SEQUENCE.
 RX MEDLINE-91107436; PubMed-2125590;
 RA Friedenham A., Fendrich G., Markl F., Gruner J.,
 RA Raschdorf F., Peter H.H.;
 RT "Duramycin B and C, two new lanthionine containing antibiotics as
 RT inhibitors of phospholipase A2. Structural revision of duramycin and
 RT cinnamycin.";
 RL J. Antibiot. 43:1403-1412(1990).
 RN [2]
 RP STRUCTURE BY NMR.
 RA Zimmermann N., Freund S., Friedenham A., Jung G.;
 RT "Solution structure of the lantibiotics duramycin B and C.";
 RL (In) Schneider C.H., Eberles A.N. (eds.);
 RL Peptides 1992, pp.519-520, ESCOM Science Publishers, Leiden (1993).
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE-93387292; PubMed-8375380;
 RA Zimmermann N., Freund S., Friedenham A., Jung G.;
 RT "Solution structures of the lantibiotics duramycin B and C.";
 RL Eur. J. Biochem. 216:419-428(1993).
 CC -1- FUNCTION: ACTS AS INHIBITOR OF PHOSPHOLIPASE A2.
 CC -1- PPM: MATURATION OF LANTHBIOTICS INVOLVES THE ENZYMIC CONVERSION OF
 CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE
 CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND
 CC CLEAVAGE OF THE MODIFIED PRECURSOR.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE B LANTHBIOTICS.
 CC Antibiotic; Lantibiotic.

FT MOD_RES 4 4 D-ALANINE.
 FT MOD_RES 6 6 D-ABU (BONDED TO 19).
 FT MOD_RES 11 11 D-ABU (AMINOBTYRIC ACID).
 FT MOD_RES 18 18 D-ABU (AMINOBTYRIC ACID).
 FT MOD_RES 19 19 LYSINOLANINE (BONDED TO 6).
 FT THIOETH 1 18 ABU-S-CYS (METHYLLANTHIONINE).
 FT THIOETH 4 14 ALA-S-CYS (LANTHIONINE).
 FT THIOETH 5 11 ABU-S-CYS (METHYLLANTHIONINE).
 SQ SEQUENCE 19 AA; 2063 MW; 0133445E27362F00 CRC64;

Query Match 19.1%; Score 21; DB 1; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 CDG 18
 DB 14 CDG 16

RESULT 16

CH60_BOVIN STANDARD: PRT: 22 AA.

AC P31081;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE 60 KDA HEAT SHOCK PROTEIN, MITOCHONDRIAL (HSP60) (60 KDA CHAPERONIN)
 DE (CPN60) (HEAT SHOCK PROTEIN 60) (HSP-60) (MITOCHONDRIAL MATRIX PROTEIN
 DE P1) (FRAGMENT).
 GN HSPD1 OR HSP60.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.

RN [1]
 RP SEQUENCE.
 RC TISSUE-ADRENAL GLAND;
 RX MEDLINE-93208180; PubMed-8096152;
 RA Dal Z., Lackland H., Stein S., Li Q., Radziewicz R., Williams R.,
 RA Sigal L.H.;
 RT "Molecular mimicry in Lyme disease: monoclonal antibody H9724 to B.
 RT burgdorferi flagellin specifically detects chaperonin-HSP60.";
 RL Biochim. Biophys. Acta 1181:97-100(1993).
 CC -1- FUNCTION: IMPLICATED IN MITOCHONDRIAL PROTEIN IMPORT AND
 CC MACROMOLECULAR ASSEMBLY. MAY FACILITATE THE CORRECT FOLDING OF
 CC IMPORTED PROTEINS. MAY ALSO PREVENT MISFOLDING AND PROMOTE THE
 CC REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED
 CC UNDER STRESS CONDITIONS IN THE MITOCHONDRIAL MATRIX.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
 CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
 DR HSSP; P06139; 1A0N.
 DR INTERPRO; IPR001844; -.
 DR PROSITE; PS00296; CHAPERONINS_CPN60; PARTIAL.
 KM Chapterone; ATP-binding; Mitochondrion.
 FT NON_TER 22
 SQ SEQUENCE 22 AA; 2303 MW; E09D2EB9934475F5 CRC64;

Query Match 19.1%; Score 21; DB 1; Length 22;
 Best Local Similarity 38.5%; Pred. No. 2.5e+03;
 Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 DYNALFHTDIL 13
 DB 9 DABLMQGVLL 21

RESULT 17

ACT7_SOYBN STANDARD: PRT: 13 AA.

AC P15987;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ACTIN 7 (FRAGMENT).
 GN SACT.
 OS Glycine max (Soybean).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 CC Fabales; Fabaceae; Papilionoideae; Glycine.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, WAYNE;
 RX MEDLINE-91346640; PubMed-2102831;
 RA Pearson L., Meagher R.B.;
 RT "Diverse soybean actin transcripts contain a large intron in the 5'
 RT untranslated leader: structural similarity to vertebrate muscle actin
 RT genes.";
 RL Plant Mol. Biol. 14:513-526(1990).

CC -1- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED
 CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBQUITOUSLY EXPRESSED
 CC IN ALL EUKARYOTIC CELLS.
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF CELL CYTOSKELETON; PLAYS AN
 CC IMPORTANT ROLE IN CYTOPLASMIC STREAMING, CELL SHAPE DETERMINATION,
 CC CELL DIVISION, ORGANELLE MOVEMENT AND EXTENSION GROWTH.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MISCELLANEOUS: THERE ARE AT LEAST 16 ACTIN GENES IN SOYBEAN.
 CC -1- SIMILARITY: BELONGS TO THE ACTIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 CC EMBL: X17120; CAA34980.1; -
 CC PIR: S15755; S15755.
 CC INTERPRO: IPR000279; -
 CC PROSITE: PS00406; ACTINS_1; PARTIAL.
 CC PROSITE: PS00432; ACTINS_2; PARTIAL.
 CC PROSITE: PS01132; ACTINS_ACT-LIKE; PARTIAL.
 CC Structural protein; Multigene family.
 CC FT NON_TER 13
 CC SEQUENCE 13 AA; 1420 MW; 8BEFF3C36D4FD05A CRC64;

Query Match 18.6%; Score 20.5; DB 1; Length 13;
 Best Local Similarity 50.0%; Pred. No. 1.8e+03;
 Matches 5; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 9 ATDLP-ACD 17
 Db 4 AEDIOPLVCD 13

RESULT 18
 ACT6_SOYBN STANDARD; PRT; 17 AA.
 AC P15966;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ACTIN 6 (FRAGMENT).
 GN SACC.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fabales; Fabaceae; Papilionoideae; Glycine.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, WAYNE;
 RX MEDLINE-91346640; PubMed-2102831;
 RA Pearson L., Mesgher R.B.;
 RT "Diverse soybean actin transcripts contain a large intron in the 5'
 RT untranslated leader: structural similarity to vertebrate muscle actin
 RT genes.";
 RL Plant Mol. Biol. 14:513-526(1990).

CC -1- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED
 CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED
 CC IN ALL EUKARYOTIC CELLS.
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF CELL CYTOSKELETON; PLAYS AN
 CC IMPORTANT ROLE IN CYTOPLASMIC STREAMING, CELL SHAPE DETERMINATION,
 CC CELL DIVISION, ORGANELLE MOVEMENT AND EXTENSION GROWTH.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MISCELLANEOUS: THERE ARE AT LEAST 16 ACTIN GENES IN SOYBEAN.
 CC -1- SIMILARITY: BELONGS TO THE ACTIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 CC EMBL: X17119; CAA34979.1; -
 CC PIR: S15754; S15754.
 CC INTERPRO: IPR000279; -
 CC PROSITE: PS00406; ACTINS_1; PARTIAL.
 CC PROSITE: PS00432; ACTINS_2; PARTIAL.

DR PROSITE: PS01132; ACTINS_ACT-LIKE; PARTIAL.
 KW Structural protein; Multigene family.
 FT* NON_TER 17
 SQ SEQUENCE 17 AA; 1749 MW; 38F4970D4BEFF3C3 CRC64;

Query Match 18.6%; Score 20.5; DB 1; Length 17;
 Best Local Similarity 50.0%; Pred. No. 2.3e+03;
 Matches 5; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 9 ATDLP-ACD 17
 Db 4 AEDIOPLVCD 13

RESULT 19
 FLAA_TREPH STANDARD; PRT; 25 AA.
 AC P21983;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-AUG-1991 (Rel. 19, Last annotation update)
 DE FLAGELLAR FILAMENT OUTER LAYER PROTEIN (SHEATH PROTEIN) (FRAGMENT).
 GN FLAA.
 OS Treponema phagedentis.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-88314903; PubMed-3045083;
 RA Norris S.J., Charon N.W., Cook R.G., Fuentes M.D., Limberger R.J.;
 RT "Antigenic relatedness and N-terminal sequence homology define two
 RT classes of periplasmic flagellar proteins of Treponema pallidum
 RT subsp. pallidum and Treponema phagedentis.";
 RL J. Bacteriol. 170:4072-4082(1988).
 CC -1- FUNCTION: COMPONENT OF THE OUTER LAYER OF THE FLAGELLA.
 CC -1- SUBUNIT: OUTER LAYER COMPOSED OF REPEATING UNITS OF FLAA AND A
 CC CORE THAT CONTAINS ONE OR ALL OF FIVE ANTIGENICALLY RELATED
 CC POLYPEPTIDE.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC FLAGELLUM.
 DR PIR: B32351; B32351.
 KW Flagella; Periplasmic.
 FT NON_TER 25
 SQ SEQUENCE 25 AA; 2676 MW; EB3841091FDBDAB3 CRC64;

Query Match 18.6%; Score 20.5; DB 1; Length 25;
 Best Local Similarity 29.4%; Pred. No. 3.4e+03;
 Matches 5; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

OY 2 VNFAFLHATDLPACDG 18
 Db 6 IDEGKLA-DIVPKNG 21

RESULT 20
 FARP_MANSE STANDARD; PRT; 10 AA.
 AC P18523;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE FMRFAMIDE-LIKE NEUROPEPTIDE.
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Sphingidae; Sphingidae; Sphinginae; Manduca.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-91045350; PubMed-2235684;
 RX Kingan T.G., Teplow D.B., Phillips J.M., Riehm J.P., Rao K.R.,
 RA Hildebrand J.G., Homberg U., Kammer A.E., Jardine I., Griffin P.R.,
 RA Hunt D.F.;
 RT "A new peptide in the FMRFamide family isolated from the CNS of the

```

RT hawkmoth, Manduca sexta.
RL Peptides 11:849-856(1990).
CC -1- FUNCTION: INCREASES THE FORCE OF NEURALLY EVOKED CONTRACTIONS IN
CC THE MAJOR POWER-PRODUCING FLIGHT MUSCLES, THE DORSAL LONGITUDINAL
CC MUSCLES AND SO IS LIKELY TO PLAY A ROLE IN SUSTAINING OR PROMOTING
CC FLIGHT BEHAVIOR PATTERNS.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR: A43977; A43977.
KW Amidation; Neuropeptide.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA: 1247 MW: D3C45229D5B1F2D2 CRC64;

Query Match
Best Local Similarity 57.1%; Score 20; DB 1; Length 10;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DYNVAF 7
   1 1 1 1
Db 2 DVVHSE 8

RESULT 21
CXAL_CONST STANDARD: PRT: 13 AA.
AC P15471;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE ALPHA-CONOTOXIN SI (S1).
OS Conus striatus (Striated cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
RN [1]
RP SEQUENCE.
RX MEDLINE-89062448; PubMed-3196703;
RA Zafaralla G.C., Ramillo C., Gray W.R., Karlstrom R., Olivera B.M.,
RA Cruz L.J.;
RT "Phylogenetic specificity of cholinergic ligands: alpha-conotoxin
RT SI. ";
RL Biochemistry 27:7102-7105(1988).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
DR PIR: A28953; A28953.
DR HSSP: P01519; INOT.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 2 7
FT DISULFID 3 13
FT MOD_RES 13 13 AMIDATION.
SQ SEQUENCE 13 AA: 1359 MW: DEE90F8FE2457EBD CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 13;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 PAC 16
   1 1 1
Db 5 PAC 7

RESULT 22
CXAX_CONGE STANDARD: PRT: 13 AA.
AC P01520;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DE ALPHA-CONOTOXIN GII.

```

```

OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
RN [1]
RP SEQUENCE.
RX MEDLINE-81191854; PubMed-7014556;
RA Gray W.R., Luque A., Olivera B.M., Barrett J., Cruz L.J.;
RT "Peptide toxins from Conus geographus venom. ";
RL J. Biol. Chem. 256:4734-4740(1981).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE-84032400; PubMed-6630187;
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT "Conotoxin M1. Disulfide bonding and conformational states. ";
RL J. Biol. Chem. 258:12247-12251(1983).
RN [3]
RP REVIEW.
RX MEDLINE-89024586; PubMed-3052286;
RA Gray W.R., Olivera B.M., Cruz L.J.;
RT "Peptide toxins from venomous conus snails. ";
RL Annu. Rev. Biochem. 57:665-700(1988).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
DR PIR: A01783; NTKN2G.
DR HSSP: P01519; INOT.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 2 7
FT DISULFID 3 13
FT MOD_RES 13 13 AMIDATION.
SQ SEQUENCE 13 AA: 1422 MW: DEEB831C39297EBD CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 13;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 PAC 16
   1 1 1
Db 5 PAC 7

RESULT 23
CXAL_CONST STANDARD: PRT: 13 AA.
AC P28878;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE ALPHA-CONOTOXIN SIA (S1A).
OS Conus striatus (Striated cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
RN [1]
RP SEQUENCE.
RX MEDLINE-91369955; PubMed-1892838;
RA Myers R.R., Zafaralla G.C., Gray W.R., Abbot J., Cruz L.J.,
RA Olivera B.M.;
RT "Alpha-conotoxins, small peptide probes of nicotinic acetylcholine
RT receptors. ";
RL Biochemistry 30:9370-9377(1991).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
DR PIR: A40312; NTKN4S.
DR Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 2 7
FT DISULFID 3 13
FT MOD_RES 13 13 AMIDATION.
SQ SEQUENCE 13 AA: 1461 MW: DEF1931982457EBD CRC64;

```

Query Match 18.2%; Score 20; DB 1; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 PAC 16
 DB 5 PAC 7

RESULT 24

CXAL_CONCN STANDARD: PRT: 14 AA.
 ID CXAL_CONCN
 AC P56973;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ALPHA-CONOTOXIN CNIA [CONTAINS: ALPHA-CONOTOXIN CNIB].
 OS Conus consors.
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 [1]
 RP SEQUENCE, SYNTHESIS, CHARACTERIZATION, AND STRUCTURE BY NMR.
 RC TISSUE-VENOM;
 RX MEDLINE-99255390; PubMed-10320362;
 RA Favreau P., Kimm I., Le Gall F., Bobenrieth M.J., Lamthanh H.,
 RA Bouet F., Servent D., Molgo J., Menez A., Letourneux Y.,
 RA Lancelin J.-M.;
 RT "Biochemical characterization and nuclear magnetic resonance
 structure of novel alpha-conotoxins isolated from the venom of Conus
 consors.";
 RT Biochemistry 38:6317-6326(1999).
 RL -I- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
 BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
 INHIBIT THEM. THIS PEPTIDE SEEMS TO BE A POTENT AND SELECTIVE
 BLOCKER OF MUSCULAR SUBTYPE OF NACHR.
 CC PDB: 1845; 09-JUL-99.
 DR Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
 KW Venom; 3D-structure.
 KM PEPTIDE 1 14 ALPHA-CONOTOXIN CNIA.
 FT PEPTIDE 3 14 ALPHA-CONOTOXIN CNIB.
 FT DISULFID 3 8
 FT DISULFID 4 14
 FT MOD.RES 14 14
 FT MOD.RES 14 14
 SQ SEQUENCE 14 AA; 1548 MW; DEE91969BF5E5BD CRC64;

Query Match 18.2%; Score 20; DB 1; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.3e+03;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 PAC 16
 DB 6 PAC 8

RESULT 25

CXAL_CONCN STANDARD: PRT: 14 AA.
 ID CXAL_CONCN
 AC P01521;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE ALPHA-CONOTOXIN MI (MI).
 OS Conus magus (Magus cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 [1]
 RP SEQUENCE.
 RX MEDLINE-83073458; PubMed-7149738;
 RA McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
 RT "Isolation and structure of a peptide toxin from the marine snail

RT Conus magus.";
 RL Arch. Biochem. Biophys. 218:329-334(1982).

RN [2]
 RP DISULFIDE BONDS.
 RX MEDLINE-84032400; PubMed-6630187;
 RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
 RT "Conotoxin MI. Disulfide bonding and conformational states.";
 RL J. Biol. Chem. 258:12247-12251(1983).

RN [3]
 RP REVIEW.
 RX MEDLINE-89024586; PubMed-3052286;
 RA Gray W.R., Olivera B.M., Cruz L.J.;
 RT "Peptide toxins from venomous Conus snails.";
 RL Ann. Rev. Biochem. 57:665-700(1988).
 CC -I- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
 BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
 INHIBIT THEM.
 DR PIR: A01784; NTKNLM.
 KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
 KM Venom.
 FT DISULFID 3 8
 FT DISULFID 4 14
 FT DISULFID 4 14
 FT MOD.RES 14 14
 FT MOD.RES 14 14
 SQ SEQUENCE 14 AA; 1499 MW; DEE91898BF5E5BD CRC64;

Query Match 18.2%; Score 20; DB 1; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.3e+03;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 PAC 16
 DB 6 PAC 8

Search completed: February 5, 2001, 10:55:56
 Job time: 496 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:47:34 : Search time 68.99 Seconds
(without alignments)
9.913 Million cell updates/sec

Title: US-08-981-824-7
Perfect score: 103
Sequence: 1 FFRVSNPATHTQDIDFLI 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

A_Geneseq_36.*
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	92.2	20	16	R72297
2	95	92.2	20	21	Y59560
3	73	70.9	14	16	R76662
4	73	70.9	14	18	W18870
5	52	50.5	20	16	R72298
6	52	50.5	20	21	Y59547
7	52	50.5	20	21	Y59589
8	35	34.0	19	13	R28755
9	35	34.0	19	16	R71643
10	34	33.0	18	21	Y85070
11	33	32.0	18	19	W44912
12	33	32.0	19	19	W44924

13	33	32.0	20	15	R56619
14	33	32.0	20	15	R61238
15	33	32.0	20	16	R67426
16	32	31.1	20	15	R56620
17	32	31.1	20	15	R61236
18	32	31.1	20	15	R61237
19	32	31.1	20	16	R67427
20	31	30.1	19	13	R26891
21	31	30.1	23	13	R26894
22	30	29.1	17	12	R13289
23	30	29.1	17	19	W87486
24	30	29.1	18	16	R63773
25	30	29.1	18	16	R66374
26	30	29.1	25	8	P70377
27	30	29.1	25	11	R00581
28	29	28.2	17	21	Y59291
29	29	28.2	20	15	R63268
30	29	28.2	20	18	W42899
31	29	28.2	25	20	Y45345
32	28	27.2	8	19	W59312
33	28	27.2	9	21	Y69593
34	28	27.2	9	21	Y50884
35	28	27.2	13	19	W87480
36	28	27.2	14	16	R71030
37	28	27.2	20	15	R74077
38	28	27.2	20	15	R74078
39	28	27.2	20	16	R67425
40	28	27.2	21	19	W65674
41	28	27.2	22	13	R28278
42	27	26.2	9	20	Y36471
43	27	26.2	10	21	Y51451
44	27	26.2	15	14	R42732
45	27	26.2	15	14	R42733

ALIGNMENTS

RESULT 1
ID R72297 standard; Peptide: 20 AA.
AC R72297;
DT 13-NOV-1995 (first entry)
XX
DE Glutamic acid decarboxylase (GAD65) fragment.
XX
KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
KW Insulin-dependent diabetes mellitus; stiff man disease.
XX
OS Homo sapiens.
XX
PN W09507992-A-
XX
PD 23-MAR-1995.
XX
PF 24-AUG-1994; 94WO-US09478.
XX
PR 17-SEP-1993; 93US-0123859.
XX
PA (REGC) UNIV CALIFORNIA.
PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
XX
DR WPI; 1995-131360/17.
XX
PT New polypeptide fragments of glutamic acid decarboxylase - for
PT diagnosis and treatment of autoimmune disease, esp. insulin
PT dependent diabetes; also related nucleic acid, vectors,
PT antihodies, hydrioma(s) etc.
XX
PS Claim 1; Page 77; 100pp; English.

CC 086481 and 086482 encode R71733 and R79105, rat and human glutamic
 CC acid decarboxylase (GAD65), respectively, from which the GAD65
 CC fragments described in R72261-R72298 were derived. These fragments
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose
 CC and treat GAD-related autoimmune disorders, such as insulin
 CC dependent diabetes mellitus or stiff man disease.
 CC
 XX Sequence 20 AA:

Query Match 92.2%; Score 95; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.2e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFRWVSNPATRHQDIDF 18
 ||||||||||||||||
 DB 3 ffrwvsnpatrhqdidf 20

RESULT 2
 Y59560 standard; peptide: 20 AA.
 Y59560:

03-APR-2000 (first entry)

GAD65 fragment, peptide #37.

GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
 insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
 therapy.

Homo sapiens.

US5998366-A.

07-DEC-1999.

09-APR-1997; 97US-0827618.

07-JUN-1995; 95US-0485725.

21-SEP-1990; 90US-0585336.

18-JUN-1991; 91US-0716909.

(REGC) UNIV CALIFORNIA.

Tobin AJ, Kaufman DL, Erlander MG;

WPI: 2000-095930/08.

Ameliorating glutamic acid decarboxylase associated autoimmune
 disorders such as insulin dependent diabetes mellitus and Stiff man
 disease -

Claim 1; Column 43; 61pp; English.

This sequence represents a fragment of the glutamic acid decarboxylase 65
 (GAD65) protein. The invention relates to a method of ameliorating GAD
 associated autoimmune disorders by administering a GAD65 peptide to the
 patient. The method can be used for ameliorating GAD associated
 autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
 and Stiff man disease. GAD65 can also be useful for screening drugs that
 alter GAD function, for generating monoclonal antibodies and in
 immunosays. GAD65 is an effective diagnostic tool for predicting IDDM
 and the diagnosis is quite easy. It is also possible to obtain much
 larger quantities of polypeptide via recombinant techniques than are
 available from natural sources.

Sequence 20 AA;

Query Match 92.2%; Score 95; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.2e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFRWVSNPATRHQDIDF 18
 ||||||||||||||||
 DB 3 ffrwvsnpatrhqdidf 20

RESULT 3

R76662 standard; peptide: 14 AA.

R76662;

05-MAR-1996 (first entry)

Peptide derived from human glutamic acid decarboxylase 21.

diabetes; T-cell subpopulation; detection; antigen production;
 diagnosis; autoimmune disease.

Homo sapiens.

DE4418091-A1.

27-JUL-1995.

24-MAY-1994; 94DE-4418091.

04-FEB-1994; 94DE-4403522.

20-JAN-1994; 94DE-4401629.

(ENDL/) ENDL J.
 (BOEF) BOEHRINGER MANNHEIM GMBH.

Albert W, Dormair K, Endl J, Jung G, Meinel E;

Stahl P, Schendel D;

WPI: 1995-264505/35.

Antigen-specific activated T-lymphocytes and their detection - by
 interaction with inventive peptide(s) of peptide-MHC complexes;
 useful in diagnosis of e.g. diabetes and autoimmune diseases

Claim 1; Fig 2; 21pp; German.

R76642-62 are derived from human glutamic acid decarboxylase and
 specifically react with T-cell subpopulations isolated from recently
 diagnosed Type-I diabetics. Pharmaceutical compns. contg. these
 peptides and those shown in R77571-72, are useful for the diagnosis of
 a disease or predispositions of immune system diseases, tumours, and
 autoimmune diseases, including diabetes. The peptides are able to detect
 specific T-cell subpopulations that are then used for antigen prochn.,
 e.g. by reinjection.

Sequence 14 AA;

Query Match 70.9%; Score 73; DB 16; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.7e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VISNPAATHQDIDF 18

DB 1 visnpatrhqdidf 14

RESULT 4

W18870 standard; peptide: 14 AA.

W18870;

XX 05-JAN-1998 (first entry)
 DT 65 kD Glutamic acid decarboxylase peptide fragment 21.
 XX
 XX
 DE
 XX
 XX
 KM GAD, 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
 KM Insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
 KM Predisposition; autoimmune; tumour; rheumatoid arthritis;
 KM multiple sclerosis.
 XX
 OS Synthetic.
 XX
 XX DE19526561-A1.
 XX
 XX 23-JAN-1997.
 PD
 XX
 XX 20-JUL-1995; 95DE-1026561.
 PF
 XX 20-JUL-1995; 95DE-1026561.
 PR
 XX 20-JUL-1995; 95DE-1026561.
 XX
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 XX
 PI Donie F, Endl J, Ganz M, Jung G, Kientsch-engel R;
 PI Pozzilli P, Stahl P;
 XX
 XX WPI; 1997-088254/09.
 DR
 XX
 XX
 PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
 PT Involving intradermal admin. of auto:reactive substances
 XX
 XX Claim 11; Fig 2; 12pp; German.
 PS
 XX
 CC W1842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predisposition to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 CC
 XX
 XX Sequence 14 AA;
 SQ

Query Match 70.9%; Score 73; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.7e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VISNPATHDIDF 18
 ||||||||
 Db 1 visnpatqdidf 14

RESULT 5
 R72298 R72298 standard; Peptide; 20 AA.
 ID
 XX
 AC R72298;
 XX
 DT 13-NOV-1995 (first entry)
 XX
 DE Glutamic acid decarboxylase (GAD65) fragment.
 XX
 KM Glutamic acid decarboxylase; GAD65; autoimmune disorders;
 KM Insulin-dependent diabetes mellitus; stiff man disease.
 XX
 OS Homo sapiens.
 XX
 XX WO9507992-A.
 PN
 XX 23-MAR-1995.
 PD

XX 24-AUG-1994; 94MO-US09478.
 PF
 XX
 PR 17-SEP-1993; 93US-0123859.
 XX
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
 PI
 XX WPI; 1995-131360/17.
 DR
 XX
 XX
 PT New polypeptide fragments of glutamic acid decarboxylase - for
 PT diagnosis and treatment of autoimmune disease, esp. insulin
 PT dependent diabetes, also related nucleic acid, vectors,
 PT antibodies, hydridoma(s) etc.
 XX
 XX
 PS Claim 1; Page 77; 100pp; English.
 XX
 XX CC 086481 and 086482 encode R71733 and R79105, rat and human glutamic
 CC acid decarboxylase (GAD65) respectively, from which the GAD65
 CC fragments described in R72261-R72298 were derived. These fragments
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose
 CC and treat GAD-related autoimmune disorders, such as insulin
 CC dependant diabetes mellitus or stiff man disease.
 CC
 XX
 XX Sequence 20 AA;
 SQ

Query Match 50.5%; Score 52; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.015; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0;

QY 11 ATHODIDFLI 20
 ||||||||
 Db 1 atqdidfll 10

RESULT 6
 Y59547
 ID Y59547 standard; peptide; 20 AA.
 XX
 AC Y59547;
 XX
 DT 03-APR-2000 (first entry)
 XX
 DE GAD65 fragment, peptide #38.
 XX
 KM GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
 KM Insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
 KM therapy.
 XX
 OS Homo sapiens.
 XX
 XX US5998366-A.
 PN
 XX
 PD 07-DEC-1999.
 XX
 PF 09-APR-1997; 97US-0827618.
 XX
 PR 07-JUN-1995; 95US-0485725.
 PR 21-SEP-1990; 90US-0586536.
 PR 18-JUN-1991; 91US-0716909.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA
 PI Tobin AJ, Kaufman DL, Erlander MG;
 PI
 XX WPI; 2000-095930/08.
 DR
 XX
 XX Ameliorating glutamic acid decarboxylase associated autoimmune
 PT disorders such as insulin dependent diabetes mellitus and Stiff man
 PT disease -
 XX

PS Claim 1; Column 33; 61pp; English.

CC This sequence represents a fragment of the glutamic acid decarboxylase 65
 CC (GAD65) protein. The invention relates to a method of ameliorating GAD
 CC associated autoimmune disorders by administering a GAD65 peptide to the
 CC patient. The method can be used for ameliorating GAD associated
 CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
 CC and Stiff man disease. GAD65 can also be useful for screening drugs that
 CC alter GAD function, for generating monoclonal antibodies and in
 CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
 CC and the diagnosis is quite easy. It is also possible to obtain much
 CC larger quantities of polypeptide via recombinant techniques than are
 CC available from natural sources.

CC Sequence 20 AA;

Query Match 50.5%; Score 52; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ATHODIDFLI 20
 |||||||
 Db 1 atgqddifli 10

RESULT 7

Y59589 standard; peptide: 20 AA.

Y59589;

03-APR-2000 (first entry)

GAD65 fragment, peptide #38.

GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
 insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
 therapy.

Homo sapiens.

US5998366-A.

07-DEC-1999.

09-APR-1997; 97US-0827618.

07-JUN-1995; 95US-0485725.

21-SEP-1990; 90US-0586536.

18-JUN-1991; 91US-0716909.

(REGC) UNIV CALIFORNIA.

Tobin AJ, Kaufman DL, Erlander MG;

WPI: 2000-095930/08.

Ameliorating glutamic acid decarboxylase associated autoimmune
 disorders such as insulin dependent diabetes mellitus and Stiff man
 disease -

Example 11; Column 43; 61pp; English.

This sequence represents a fragment of the glutamic acid decarboxylase 65
 (GAD65) protein. The invention relates to a method of ameliorating GAD
 associated autoimmune disorders by administering a GAD65 peptide to the
 patient. The method can be used for ameliorating GAD associated
 autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
 and Stiff man disease. GAD65 can also be useful for screening drugs that
 alter GAD function, for generating monoclonal antibodies and in
 immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
 and the diagnosis is quite easy. It is also possible to obtain much

CC larger quantities of polypeptide via recombinant techniques than are
 CC available from natural sources.

CC Sequence 20 AA;

Query Match 50.5%; Score 52; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ATHODIDFLI 20
 |||||||
 Db 1 atgqddifli 10

RESULT 8

R28755 standard; Protein: 19 AA.

R28755;

20-APR-1993 (first entry)

Synthetic rabbit GAD antibody 1266 raising peptide.

Glutamic acid decarboxylase; rabbit; immunoreactive.

Synthetic.

WO9220811-A.

26-NOV-1992.

14-MAY-1992; 92WO-US04079.

15-MAY-1991; 91US-0702162.

(UNIM) UNIV WASHINGTON.

(ZYMO) ZYMOGENETICS INC.

Foster DC, Grubin CE, Hagopian W, Karlsen AE, Lernmark A, O'Hara PJ;

WPI: 1992-415789/50.

Polynucleotide encoding human islet glutamic acid decarboxylase -
 used to test for auto-antibodies against itself to diagnose
 insulin dependent diabetes mellitus

Example; Page 33; 45pp; English.

The peptide was synthesized based on the C-terminal sequence of
 human islet glutamic acid decarboxylase (GAD). It was used to raise
 antibody 1266 in rabbits, the antibody was used to test transfectants
 for production of human GAD by immunocytochemistry.

Sequence 19 AA;

Query Match 34.0%; Score 35; DB 13; Length 19;

Best Local Similarity 77.8%; Pred. No. 11;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 12 THODIDFLI 20
 |||||||

Db 1 tgsddifli 9

RESULT 9

R71643 standard; Peptide; 19 AA.

R71643;

DT 21-SEP-1995 (first entry)
 XX Human GAD C-terminal peptide.
 DE
 XX GAD; glutamic-acid-decarboxylase; glutamate-decarboxylase;
 KW non-insulin-dependent diabetes; NIDDM; diagnosis; autoantibody;
 KW pancreas; islet.
 XX
 OS Homo sapiens.
 XX
 PN W09507464-A.
 XX
 PD 16-MAR-1995.
 XX
 PF 02-SEP-1994; 94WO-US09937.
 XX
 PR 07-SEP-1993; 93US-0117907.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Hagopian W, Karlisen AE, Landin-Olsson M, Lernmark A;
 DR WPI: 1995-123512/16.
 XX
 XX Predicting the clinical course of diabetes in patients with
 PT non-insulin dependent diabetes mellitus - by detecting the
 PT presence of autoantibodies to human islet cell glutamic acid
 PT decarboxylase.
 XX
 PS Example II: Page 26; 62pp; English.
 XX
 CC Human islet cDNA (Q86046) was expressed in tk- ts13 BHK cells
 CC (ATCC CRL 1632). Transfectants were selected in methotrexate
 CC medium. The immunoreactivity of the human GAD protein obtained
 CC was tested using antibodies raised in rabbits against a C-terminal
 CC peptide (R71643) of human GAD.
 CC
 XX
 SQ Sequence 19 AA:
 Query Match 34.0%; Score 35; DB 16; Length 19;
 Best Local Similarity 77.8%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 12 THQDIDFLI 20
 Db 1 tgsdldfl 9
 RESULT 10
 Y85070 Y85070 standard; peptide; 18 AA.
 XX
 AC Y85070;
 XX
 DT 20-JUN-2000 (first entry)
 XX
 DE Immunogenic peptide #8 derived from FMDV non-structural protein 3C.
 XX
 KW Foot and mouth disease; FMDV; epitope; anti-FMDV antibody; diagnosis;
 KW vaccine; immunoreagent; non-structural protein.
 XX
 OS Foot and mouth disease virus.
 XX
 PN US6048538-A.
 XX
 PD 11-APR-2000.
 XX
 PF 03-OCT-1997; 97US-0943173.
 XX
 PR 03-OCT-1997; 97US-0943173.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.

XX
 PI Y1 Wang C, De Chen P, Shen F;
 XX
 DR WPI: 2000-328034/28.
 XX
 XX Peptide from non-structural proteins useful for detecting
 PT foot-and-mouth disease virus infection in animal comprises an
 PT immunogenic epitope specific to antibodies
 XX
 PS Claim 1; Column 25-26; 23pp; English.
 XX
 CC This sequence represents a peptide derived from the non-structural
 CC protein 3C of the foot and mouth disease virus (FMDV). The amino acid
 CC sequence of the peptide corresponds to an immunodominant region of the
 CC FMDV non-structural protein and is an immunogenic epitope specific for
 CC anti-FMDV antibodies. The invention relates to peptides derived from
 CC three non-structural FMDV proteins. The peptides are useful as
 CC immunoreagents i.e. reagent in an immunoassay in detecting antibodies to
 CC FMDV or diagnosis of FMDV infection. The immunoassays are used to screen
 CC body fluids and tissues for the presence of FMDV-reactive antibodies and
 CC aid in the diagnosis of FMDV infection and differentiate infected from
 CC the vaccinated animals.
 CC
 XX
 SQ Sequence 18 AA:
 Query Match 33.0%; Score 34; DB 21; Length 18;
 Best Local Similarity 42.9%; Pred. No. 16;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 3 RMVTSNPATRHQDI 16
 Db 3 rllfsgaelytkd 16
 RESULT 11
 W44912 W44912 standard; peptide; 18 AA.
 XX
 AC W44912;
 XX
 DT 29-SEP-1998 (first entry)
 XX
 DE Spacer peptide for chimeric glycoprotein AMO-Delta-Pro2.
 XX
 KW Chimeric; polypeptide; beta-turn helix; facilitation; inhibition;
 KW retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
 KW myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
 XX
 OS Synthetic.
 XX
 PN W09744474-A2.
 XX
 PD 27-NOV-1997.
 XX
 PF 16-MAY-1997; 97WO-FR00870.
 XX
 PR 20-MAY-1996; 96FR-0006234.
 XX
 PA (CNRS) CENT NAT RECH SCI.
 XX
 PI Cosset FL, Russell SJ, Valsesia S;
 XX
 DR WPI: 1998-018526/02.
 XX
 PT Introducing genes into eukaryotic cells using peptide with two
 PT receptor binding regions - especially as part of a viral
 PT glyco-protein, used in gene therapy to ensure precise targeting of
 PT vectors to particular cell types
 XX
 PS Example 2; Page 40; 117pp; French.
 XX
 CC A novel chimeric protein for gene transfer comprises, at the N-terminus,

CC a protein binding a target molecule on the cell surface, and, at the
CC C-terminus, a second protein recognising a secondary target molecule on
CC the cell surface. The two domains are separated by a peptide comprising
CC at least 30% Pro residues arranged so as to induce approximately 180 deg.
CC folds (beta or reverse turns) in the peptide sequence, the folds being
CC regularly spread to form a polyprotein beta-turn helix structure. The
CC whole construct especially forms part of a (retro)viral envelope
CC glycoprotein. A series of chimeric envelope glycoproteins were
CC constructed and were titrated, against cells expressing either the Ram-1
CC and/or Rec-1 cell surface receptors. This sequence is the Pro-rich
CC spacer peptide derived from the AHO-Delta-Pro2 envelope construct. The
CC pro-rich peptide acts by facilitating or inhibiting binding between the
CC C-terminal protein and its receptor, and inhibition occurs as long as the
CC N-terminal protein does not interact with its target. The recombinant
CC (retro)viral particles are used for in vitro or ex vivo transfer of
CC nucleic acid into target cells, especially for gene therapy, e.g. in
CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC human immunodeficiency virus) infections, etc.

Query Match	32.0%	Score	33	DB	19	Length	18
Best Local Similarity	85.7%	Pred	No	24			
Matches	6	Conservative	0	Mismatches	1	Indels	0
						Gaps	0

QY	8	NP	A	T	H	Q	14
		1	1	1	1	1	
Db	11	np	a	p	h	q	17

Db 11 nřapnq 17

RESULT 12

ID W44924 standard; peptide; 19 AA.

AC W44924;

DT 29-SEP-1998 (first entry)

DE Pro-rich spacer for chimeric envelope glycoprotein DeltaPro2.

KM Chimeric; polypyrrolone; beta-turn helix; facilitation; inhibition;
KM retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KM myopathy; lysosomal disorder; cancer; human immunodeficiency virus.

OS Synthetic.

PN W09744474-A2.

PD 27-NOV-1997

PF 16-MAY-1997; 97WO-FR00870.

PR 20-MAY-1996; 96FR-0006234.

PA (CNRS) CENT NAT RECH SCI.

PI Cosset FL, Russell SJ, Valsesia S;

DR WPI; 1998-018526/02.

PT Introducing genes into eukaryotic cells using peptide with two
PT receptor binding regions - especially as part of a viral
PT glyco-protein, used in gene therapy to ensure precise targeting of
PT vectors to particular cell types

PS Example 3; Page 42; 117pp; French

CC A novel chimeric protein for gene transfer compriess¹, at the N-terminus
CC a protein binding a target molecule on the cell surface, and, at the
CC C-terminus, a second protein recognising a secondary target molecule on
CC the cell surface. The two domains are separated by a peptide compising
CC at least 30% Pro residues arranged so as to induce approximately 180 deg

CC folds (beta or reverse turns) in the peptide sequence, the folds being
CC regularly spread to form a polyproline beta-turn helix structure. The
CC whole construct especially forms part of a (retro)viral envelope
CC glycoprotein. A series of chimeric envelope glycoproteins were
CC constructed and targeted to either the epidermal growth factor receptor
CC (EGFR) or the major histocompatibility complex (MHC) class I receptor.
CC Peptides W44923-W44927 are examples of the spacer peptides used in the
CC construction of the chimeric envelope glycoproteins. The pro-rich peptide
CC acts by facilitating or inhibiting binding between the C-terminal protein
CC and its receptor, and inhibition occurs as long as the N-terminal protein
CC does not interact with its target. The recombinant (retro)viral particles
CC are used for in vitro or ex vivo transfer of nucleic acid into target
CC cells, especially for gene therapy, e.g. in cystic fibrosis, myopathy,
CC lysosomal disorders, cancer, viral (especially human immunodeficiency
CC virus) infections, etc.

SQ Sequence 19 AA:

Query Match	32.08%	Score 33	DB 19	Length 19
Best Local Similarity	85.78%	Pred. No.	25	
Matches	6	Conservative	0	Mismatches 1; Indels 0; Gaps 0

```
QY      8 NPAATHQ 14
         |||||
Db      12 npaaphq 18
```

Db 12 npaphq 18

RESULT 13

ID	standard; peptide; 20 AA.
R56619	

AC R56619;

DT 24-MAR-1995 (first entry)

DE Alginic acid Al-III-2 lyase N-terminal sequence

KW Alginic acid lyase; Al-II-2; 4-deoxy-5-ketonic acid;
KW Alginic acid lyase; Al-II-2; 4-deoxy-5-ketonic acid;
KW Alginic acid lyase; Al-II-2; 4-deoxy-5-ketonic acid;

XX
XX
[E]architectural on OMC-6 (FEBRU P-12159)

XX	XX	Location/Qualifiers
EH	Yov	

ET	
EJ	Misc-difference 14
ET	/note= "not defined"

XX
PN .JP06197760-A

AA 19-JUL-1994
PD

28-DEC-1992

28-DEC-1992: 92JTP-0348465

AA
PA (GNZE) GUNZE KK.

XX
FA
XX

(ZKAB / OIBOM KAGHANU LKHODIN AN.

WFL; 1554-2000/0/33;
DN
XX

PT then 4-deoxy-5-keto
F1 Arginine acin lyase(

PT then 4deoxy-5-keto:ironic acid - is used for treatment of
PT pulmonary cystic fibrosis
XX
PS Claim 1; Page 9; 11pp. Japanese.

CC Alginate acid Al-II-2 Lyase has t

CC Alginic acid Δ -II-2 lyase has the N-terminal sequence R56619 and
CC decomposes alginic acid into a sugar having a double bond between C4
CC and C5 and finally into 4-deoxy-5-ketouronic acid. The enzyme has
CC optimal pH 8.0, optimal temp. 70 deg.C and mol.wt. 25000 and is
CC useful for treating pulmonary cystic fibrosis caused by *Pseudomonas*
CC *aeruginosa*.

SQ Sequence 20 AA:

Query Match 32.0%; Score 33; DB 15; Length 20;

Best Local Similarity 66.7%; Pred. No. 27; Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 9 PAATHODID 17
||| |
Db 2 paaahsids 10

RESULT 14

R61238 R61238 standard; Protein; 20 AA.

XX AC R61238;

XX DT 27-APR-1995 (first entry)

XX DE Alginate lyase II N-terminal sequence.

XX KM N-terminal; alginate lyase; Al-I; Al-III; enzyme; alginic acid;

XX KW cystic fibrosis.

XX OS Pseudomonas sp.

XX FH Key Location/Qualifiers
FT Misc-difference 14 /label= Undefined amino acid

XX FT

XX PN WO9419006-A.

XX PD 01-SEP-1994.

XX PF 24-FEB-1993; 93WO-JP00227.

XX PR 24-FEB-1993; 93WO-JP00227.

XX PA (GNZE) GUNZE LTD.

XX PA (SAKB) OTSUKA KAGAKU KK.

XX PI Abe S, Hisano T, Kimura A, Murata K, Nishimura M;

XX PI Okayama K, Yamaguchi H, Yamashita T, Yonemoto Y;

XX DR WPI; 1994-293999/36.

XX PT New alginate lyase from Pseudomonas sp. - useful for the

XX PT treatment of cystic fibrosis

XX PS Disclosure; Page 25; 36pp; Japanese.

XX CC This sequences represents the N-terminal fragment of alginate lyase II.

XX CC Alginate lyase I and III (Al-I and Al-III) have molecular weights of 60

XX CC and 38 kD. The alginate lyases break down alginic acid or viscous

XX CC substances formed by cystic fibrosis sufferers.

XX SQ Sequence 20 AA;

Query Match 32.0%; Score 33; DB 15; Length 20;

Best Local Similarity 66.7%; Pred. No. 27; Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 9 PAATHODID 17
||| |
Db 2 paaahsids 10

RESULT 15

R67426 R67426 standard; Peptide; 20 AA.

XX ID R67426 standard; Peptide; 20 AA.

AC R67426;

XX DT 31-AUG-1995 (first entry)

XX DE N-terminal fragment of alginic acid lyase. Al-II-2.

XX KM alginic acid lyase; flavobacterium; recombinant production;

XX KW N-terminal fragment.

XX OS Flavobacterium sp.

XX FH Key Location/Qualifiers
FT Misc-difference 14 /note= "unknown, not identified in the specification"

XX FT

XX PN JP06319569-A.

XX PD 22-NOV-1994.

XX PF 14-MAY-1993; 93JP-0113149.

XX PR 14-MAY-1993; 93JP-0113149.

XX PA (GNZE) GUNZE KK.

XX PA (SAKB) OTSUKA KAGAKU YAKUHIN KK.

XX DR WPI; 1995-040321/06.

XX PT alginic acid lyase expression gene - used in the prepn. of

XX PT alginic acid lyase

XX PS Disclosure; Page 3; 16pp; Japanese.

XX CC R67425-27 are the N-terminal fragments of alginic acid lyase (Al) -I,

XX CC -II-2 and -III respectively. The amino acid sequence of alginic acid

XX CC lyase is shown in R67424 and is encoded by the gene shown in Q75342.

XX CC It was isolated from Flavobacterium sp. by standard genetic techniques.

XX CC It is used to produce the recombinant enzyme in large quantities.

XX SQ Sequence 20 AA;

Query Match 32.0%; Score 33; DB 16; Length 20;

Best Local Similarity 66.7%; Pred. No. 27; Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 9 PAATHODID 17
||| |
Db 2 paaahsids 10

RESULT 16

R56620 R56620 standard; peptide; 20 AA.

XX AC R56620;

XX DT 24-MAR-1995 (first entry)

XX DE alginic acid Al-III lyase N-terminal sequence.

XX KW alginic acid lyase; Al-III; 4-deoxy-5-ketonic acid;

XX KM pulmonary cystic fibrosis; Pseudomonas aeruginosa.

XX OS Flavobacterium sp. ORC-6 (FERM P-12159).

XX PN JP06197760-A.

XX PD 19-JUL-1994.

XX PF 28-DEC-1992; 92JP-0348465.

XX PR 28-DEC-1992; 92JP-0348465.

XX (GNZE) GUNZE KK.
 PA (SAKB) OTSUKA KAGAKU YAKUHIN KK.
 XX
 DR WPI: 1994-268676/33.
 XX
 PT Alginic acid lyase(s) which decompose alginic acid into sugar and
 XX then 4-deoxy-5-keto:uronic acid - is used for treatment of
 PT pulmonary cystic fibrosis
 XX
 PS Claim 2; Page 9; 11pp; Japanese.
 XX
 CC Alginic acid Al-III lyase has the N-terminal sequence R56620 and
 CC decomposes alginic acid into a sugar having a double bond between C4
 CC and C5 and finally into 4-deoxy-5-ketouronic acid. The enzyme has
 CC optimal pH 8.0, optimal temp. 70 deg C and mol.wt. 38000 and is
 CC useful for treating pulmonary cystic fibrosis caused by Pseudomonas
 CC aeruginosa.
 CC
 SQ Sequence 20 AA;
 XX
 Query Match 31.1%; Score 32; DB 15; Length 20;
 Best Local Similarity 31.2%; Pred. No. 39;
 Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 OY 1 FFRWISNPAATHODI 16
 I : I : I : I : I : I :
 DB 3 fdgavvkdphtasyvdv 18
 XX
 RESULT 17
 R61236
 ID R61236 standard; peptide; 20 AA.
 XX
 AC R61236;
 XX
 DT 27-APR-1995 (first entry)
 XX
 DE Alginic acid lyase I N-terminal sequence.
 XX
 KM N-terminal; alginic acid; Al-I; Al-III; enzyme; alginic acid;
 KM cystic fibrosis.
 XX
 OS Pseudomonas sp.
 XX
 PN WO9419006-A.
 XX
 PD 01-SEP-1994.
 XX
 PF 24-FEB-1993; 93WO-JP00227.
 XX
 PR 24-FEB-1993; 93WO-JP00227.
 XX
 PA (GNZE) GUNZE LTD.
 PA (SAKB) OTSUKA KAGAKU KK.
 XX
 PI Abe S, Hisano T, Kimura A, Murata K, Nishimura M;
 PI Okayama K, Yamaguchi H, Yamashita T, Yonemoto Y;
 XX
 DR WPI: 1994-293999/36.
 XX
 PT New alginic acid lyase from Pseudomonas sp. - useful for the
 PT treatment of cystic fibrosis
 XX
 PS Claim 2; Page 24; 36pp; Japanese.
 XX
 CC The sequences given in R61236-37 represent the N-terminal fragments
 CC of alginic acid lyase I and III (Al-I and Al-III), respectively. These
 CC enzymes have molecular weights of 60 and 38 kD. The alginic acid
 CC break down alginic acid or viscous substances formed by
 CC cystic fibrosis sufferers.
 CC

SQ Sequence 20 AA;
 XX
 Query Match 31.1%; Score 32; DB 15; Length 20;
 Best Local Similarity 31.2%; Pred. No. 39;
 Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 OY 1 FFRWISNPAATHODI 16
 I : I : I : I : I : I :
 DB 3 fdgavvkdphtasyvdv 18
 XX
 RESULT 18
 R61237
 ID R61237 standard; peptide; 20 AA.
 XX
 AC R61237;
 XX
 DT 27-APR-1995 (first entry)
 XX
 DE Alginic acid lyase III N-terminal sequence.
 XX
 KM N-terminal; alginic acid; Al-I; Al-III; enzyme; alginic acid;
 KM cystic fibrosis.
 XX
 OS Pseudomonas sp.
 XX
 PN WO9419006-A.
 XX
 PD 01-SEP-1994.
 XX
 PF 24-FEB-1993; 93WO-JP00227.
 XX
 PR 24-FEB-1993; 93WO-JP00227.
 XX
 PA (GNZE) GUNZE LTD.
 PA (SAKB) OTSUKA KAGAKU KK.
 XX
 PI Abe S, Hisano T, Kimura A, Murata K, Nishimura M;
 PI Okayama K, Yamaguchi H, Yamashita T, Yonemoto Y;
 XX
 DR WPI: 1994-293999/36.
 XX
 PT New alginic acid lyase from Pseudomonas sp. - useful for the
 PT treatment of cystic fibrosis
 XX
 PS Claim 4; Page 24-25; 36pp; Japanese.
 XX
 CC The sequences given in R61236-37 represent the N-terminal fragments
 CC of alginic acid lyase I and III (Al-I and Al-III), respectively. These
 CC enzymes have molecular weights of 60 and 38 kD. The alginic acid
 CC break down alginic acid or viscous substances formed by
 CC cystic fibrosis sufferers.
 CC
 SQ Sequence 20 AA;
 XX
 Query Match 31.1%; Score 32; DB 15; Length 20;
 Best Local Similarity 31.2%; Pred. No. 39;
 Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 OY 1 FFRWISNPAATHODI 16
 I : I : I : I : I : I :
 DB 3 fdgavvkdphtasyvdv 18
 XX
 RESULT 19
 R67427
 ID R67427 standard; peptide; 20 AA.
 XX
 AC R67427;
 XX
 DT 31-AUG-1995 (first entry)

DE N-terminal fragment of alginic acid lyase, Al-III.
XX
XX alginic acid lyase; flavobacterium; recombinant production;
KW N-terminal fragment.
XX
XX Flavobacterium sp.
OS
XX JP06319569-A.
PN
XX
XX 22-NOV-1994.
PD
XX
XX 14-MAY-1993; 93JP-0113149.
PE
XX
XX 14-MAY-1993; 93JP-0113149.
PR
XX
XX
PA (GNZE) GUNZE KK.
PA (SAKB) OTSUKA KAGAKU YAKUHIN KK.
XX
XX WPI: 1995-040321/06.
DR
XX
XX
PT Alginic acid lyase expression gene - used in the prepn. of
PT alginic acid lyase
XX
XX
XX
PS Disclosure: Page 3; 16pp; Japanese.
XX
XX R67425-27 are the N-terminal fragments of alginic acid lyase (Al)-I,
CC -II-2 and -III respectively. The amino acid sequence of alginic acid
CC lyase is shown in R67424 and is encoded by the gene shown in 075342.
CC It was isolated from Flavobacterium sp. by standard genetic techniques
CC It is used to produce the recombinant enzyme in large quantities.
XX
XX
SQ Sequence 20 AA:

Query Match	31.1%	Score 33	DB 16	Length 20
Best Local Similarity	31.2%	Pred. No. 39		
Matches	5	Conservative	6	Mismatches 5
				Indels 0
				Gaps 0
QY	1	FFRWYISNPATHTODI	16	
	1	1	1	1
Db	3	fdgavvkptasyadv	18	

PT Hybrid polio virus useful as vaccine against HIV-1 infections -
PT contains epitope of heterologous protein inserted into the BC
PT loop of polio virus
XX
PS Disclosure; Page 22; 35pp; English.
XX
CC The sequences given in R26879-93 are portions of the HIV-1 principle
CC neutralising domain (PND) protein which were used as neutralising
CC epitopes. HIV-1 PND is an approx. 40 amino acid region of the
CC external envelope protein gp120 which forms a looped structure in
CC native gp120. The epitopes were inserted into the BC loop of a hybrid
CC poliovirus. This construct could be used as a vaccine. The vaccine
CC may be used for prophylaxis or treatment of human patients infected
CC with HIV-1.
XX
SQ Sequence 19 AA;

Query Match	30.1%	Score 31:	DB 13:	Length 19:
Similarity	55.6%	Pred. NC:	55:	
Local		Matches	5:	Conservative
			2:	Mismatches
				Indels
				Gaps
				0

```
QY      4 MVISNPAAT 12
        | : ||| : |
Db      5 mtvdnpast 13
```


KW major outer membrane protein; picornavirus; poliovirus; PVI;
 XX diagnostic; therapeutic; capsid protein.
 XX Chlamydia trachomatis serovar A.
 OS MO9426900-A.
 XX
 XX
 XX 24-NOV-1994.
 PD
 XX
 XX 12-MAY-1994; 94WO-CA00262.
 PF
 XX
 XX 13-MAY-1993; 93US-0060978.
 PR
 XX
 XX (CONN-) CONNAUGHT LAB LTD.
 PA
 XX
 XX Caldwell HD, Klein MH, Murdin AD, Oomen RP;
 PI
 XX WPI: 1995-006796/01.
 DR
 XX
 XX
 XX New hybrid picornaviruses expressing chlamydial epitopes - used
 PT to develop prods. for vaccination, diagnosis, treatment of
 PT chlamydial infections and prodn. of immunological reagents
 XX
 XX
 PS Disclosure: Page 15; 99pp; English.
 CC
 XX To obtain a poliovirus-1 (PVI)/chlamydial hybrid, the PVI
 CC Mahoney CDNA clone pT7XLD was modified to encode epitopes from
 CC C. trachomatis ser. A MOMP VDI. The mutagenesis cartridge was
 CC contained between PV nucleotides 2753-91 (given in 079003) which
 CC encode PV amino acids 1092-1104 (R66375) which include the BC loop
 CC of capsid protein VP1 (R66386). The polio-specific DNA within the
 CC cartridge, encoding the sequence given in R66369, was replaced by
 CC oligonucleotides encoding VDI peptides R66370-74. Viable PV
 CC expressing chlamydial immunostimulatory epitope was expressed in
 CC Vero cells.
 CC
 XX
 XX Sequence 18 AA;
 SQ

Query Match 29.1%; Score 30; DB 16; Length 18;
 Best Local Similarity 55.6%; Pred. No. 77;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 NPAATHODI 16
 |||:|
 |||:|
 DB 1 npasttsdv 9

RESULT 25
 R66374
 ID R66374 standard; peptide: 18 AA.
 XX
 AC R66374;
 XX
 DT 27-JUN-1995 (first entry)
 XX
 DE Chlamydial MOMP VDI epitope.
 XX
 XX Chlamydial epitope; variable domain; MOMP;
 KM major outer membrane protein; picornavirus; poliovirus; PVI;
 KM diagnostic; therapeutic; capsid protein.
 XX
 OS Chlamydia trachomatis serovar A.
 XX
 XX MO9426900-A.
 PN
 XX
 XX 24-NOV-1994.
 PD
 XX
 XX 12-MAY-1994; 94WO-CA00262.
 PF
 XX
 XX 13-MAY-1993; 93US-0060978.
 PR
 XX
 XX (CONN-) CONNAUGHT LAB LTD.

XX
 PI Caldwell HD, Klein MH, Murdin AD, Oomen RP;
 XX
 DR WPI: 1995-006796/01.
 XX
 XX
 XX New hybrid picornaviruses expressing chlamydial epitopes - used
 PT to develop prods. for vaccination, diagnosis, treatment of
 PT chlamydial infections and prodn. of immunological reagents
 XX
 XX
 PS Disclosure: Page 15; 99pp; English.
 CC
 XX To obtain a poliovirus-1 (PVI)/chlamydial hybrid, the PVI
 CC Mahoney CDNA clone pT7XLD was modified to encode epitopes from
 CC C. trachomatis ser. A MOMP VDI. The mutagenesis cartridge was
 CC contained between PV nucleotides 2753-91 (given in 079003) which
 CC encode PV amino acids 1092-1104 (R66375) which include the BC loop
 CC of capsid protein VP1 (R66386). The polio-specific DNA within the
 CC cartridge, encoding the sequence given in R66369, was replaced by
 CC oligonucleotides encoding VDI peptides R66370-74. Viable PV
 CC expressing chlamydial immunostimulatory epitope was expressed in
 CC Vero cells.
 CC
 XX
 XX Sequence 18 AA;
 SQ

Query Match 29.1%; Score 30; DB 16; Length 18;
 Best Local Similarity 55.6%; Pred. No. 77;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 NPAATHODI 16
 |||:|
 |||:|
 DB 1 npasttsdv 9

Search completed: February 5, 2001, 10:47:36
 Job time: 630 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:40:11 ; Search time 93.77 Seconds
(without alignments)
24.999 Million cell updates/sec

Title: US-08-981-824-1

Sequence: 1 DVNYAFLLHATDLPACDGER 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30.5	27.7	22	12_085712	085712 rous sarcom
2	30	27.3	23	8_09TJ56	09TJ56 abies fabri
3	30	27.3	23	8_09TJ54	09TJ54 abies farge
4	30	27.3	23	8_09TJ52	09TJ52 abies filma
5	30	27.3	23	8_09TJ50	09TJ50 abies frase
6	30	27.3	23	8_09TJ48	09TJ48 abies holop
7	30	27.3	23	8_09TJ46	09TJ46 abies homol
8	30	27.3	23	8_09TJ44	09TJ44 abies korea
9	30	27.3	23	8_09TJ42	09TJ42 abies lasio
10	30	27.3	23	8_09TJ40	09TJ40 abies marie
11	30	27.3	23	8_09TJ38	09TJ38 abies neph
12	30	27.3	23	8_09TJ36	09TJ36 abies sach
13	30	27.3	23	8_09TJ34	09TJ34 abies sibir
14	30	27.3	23	8_09TJ32	09TJ32 abies veit
15	30	27.3	25	2_007916	007916 mycobacteri
16	29	26.4	20	4_015988	015988 homo sapien
17	29	26.4	22	2_034195	034195 ehrlichia r
18	28	25.5	15	12_085713	085713 rous sarcom
19	28	25.5	21	10_0958J0	0958J0 bryonia dlo

20	28	25.5	23	6_09TRC6	09TRC6 cants fam1
21	28	25.5	24	5_094377	094377 caenorhabd1
22	27	24.5	21	6_09TRU55	09TRU55 cebus apell
23	26	23.6	18	7_030216	030216 homo sapien
24	26	23.6	19	4_09UMM9	09umm9 homo sapien
25	26	23.6	20	11_09OV03	09ov03 rattus sp.
26	26	23.6	21	11_088229	088229 mus muschul
27	26	23.6	21	11_09RIH1	09rih1 rattus norv
28	26	23.6	24	2_052388	052388 anabaena f1
29	26	23.6	24	5_094369	094369 caenorhabd1
30	26	23.6	24	5_094374	094374 caenorhabd1
31	25.5	23.2	20	5_09TWN0	09twn0 macrobelle
32	25.5	23.2	22	8_09TJ27	09tj27 pisum sativ
33	25	22.7	14	10_09SAP8	09sap8 pisum sativ
34	25	22.7	18	4_014009	014009 homo sapien
35	25	22.7	20	4_09UCA0	09uca0 homo sapien
36	25	22.7	21	11_09OV03	09ov03 rattus sp.
37	25	22.7	22	4_09MS53	09ms53 homo sapien
38	25	22.7	24	5_094373	094373 caenorhabd1
39	25	22.7	24	6_09TS01	09ts01 bos taurus
40	24	21.8	13	12_09WMS5	09wms5 sigma virus
41	24	21.8	15	6_09MYT7	09myt7 sus scrofa
42	24	21.8	19	2_031045	031045 streptomyce
43	24	21.8	19	8_099724	099724 dryocopus
44	24	21.8	19	8_099728	099728 buteo jama1
45	24	21.8	19	13_P87468	P87468 xenopus lae

ALIGNMENTS

RESULT 1					
ID 085712	PRELIMINARY:	PRT:	22 AA.		
AC 085712:					
DT 01-NOV-1996 (TREMBLrel. 01, Created)					
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)					
DE (BRYAN HIGH-TITER STRAIN) SRC (FRAGMENT).					
OS Rous sarcoma virus.					
OC Viruses; Retroid viruses; Retroviridae; Avian type C retroviruses.					
NCBI_Taxid=11886;					
RN [1]					
RP SEQUENCE FROM N.A.					
RX MEDLINE=84115080; PubMed=6319754;					
RA Lerner T.L.; Hanafusa H.;					
RT "DNA sequence of the Bryan high-titer strain of Rous sarcoma virus: extent of env deletion and possible genealogical relationship with other viral strains."					
RL J. VIROL. 49:549-556(1984).					
DR EMBL; K03366; AAA42559.1; -.					
FT NON_TER					
FT NON_TER					
SQ SEQUENCE	22 AA;	2519 MW;	EOBB3BB6F1DE4F3F CRC64;		
Query Match	27.7%;	Score 30.5;	DB 12;	Length 22;	
Best Local Similarity	53.8%;	Pred. No. 2e+02;			
Matches	7;	Conservative	2;	Mismatches	3;
				Indels	1;
				Gaps	1;
OY 4 YAFLLHATDLPAC 16					
DB 5 FEYLQAQLLPAC 16					
RESULT 2					
ID 09TJ56	PRELIMINARY:	PRT:	23 AA.		
AC 09TJ56:					
DT 01-MAY-2000 (TREMBLrel. 13, Created)					
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)					
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)					
DE P700 APOPROTEIN IB (FRAGMENT).					
GN PSAB.					

```

OS Abies fabri.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=97166;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsumura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029670; BAA87992.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 23 AA; 2485 MW; C19201F5B9BE67C8 CRC64;

Query Match
Best Local Similarity 27.3%; Score 30; DB 8; Length 23;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 2 VNVAFLHATDLPACDGE 19
DB 4 VGIFTYAFLIASTSGK 21

RESULT 3
ID 09TJ54 PRELIMINARY; PRT; 23 AA.
AC 09TJ54:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies fargesii.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=97167;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsumura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029671; BAA87994.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 23 AA; 2485 MW; C19201F5B9BE67C8 CRC64;

Query Match
Best Local Similarity 27.3%; Score 30; DB 8; Length 23;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 2 VNVAFLHATDLPACDGE 19
DB 4 VGIFTYAFLIASTSGK 21

RESULT 4
ID 09TJ52 PRELIMINARY; PRT; 23 AA.
AC 09TJ52:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies firma.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

```

```

OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=78260;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsumura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029672; BAA87996.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 23 AA; 2485 MW; C19201F5B9BE67C8 CRC64;

Query Match
Best Local Similarity 27.3%; Score 30; DB 8; Length 23;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 2 VNVAFLHATDLPACDGE 19
DB 4 VGIFTYAFLIASTSGK 21

RESULT 5
ID 09TJ50 PRELIMINARY; PRT; 23 AA.
AC 09TJ50:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies fraseri.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=97172;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsumura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029673; BAA87998.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 23 AA; 2485 MW; C19201F5B9BE67C8 CRC64;

Query Match
Best Local Similarity 27.3%; Score 30; DB 8; Length 23;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 2 VNVAFLHATDLPACDGE 19
DB 4 VGIFTYAFLIASTSGK 21

RESULT 6
ID 09TJ48 PRELIMINARY; PRT; 23 AA.
AC 09TJ48:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies holophylla.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=97168;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Suyama Y., Tsunmura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029674; BAA88000.1; -.
KW Chloroplast.
FT NON TER
SO SEQUENCE 23 AA; 2485 MW; C19201F5B9BE67C8 CRC64;

Query Match 27.3%; Score 30; DB 8; Length 23;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 VNVAFLHATDLLPACDGE 19
| | | | | : : | :
| | | | | : : | :
Db 4 VGIIFTVA AFLIASTSGK 21

RESULT 7
ID Q9TJ46 PRELIMINARY; PRT; 23 AA.
AC Q9TJ46;
DT 01-MAY-2000 (TREMBLrel. 13; Created)
DT 01-MAY-2000 (TREMBLrel. 13; Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13; Last annotation update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies homolepis.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBL_taxonomy=76261;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsunmura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029675; BAA88002.1; -.
KW Chloroplast.
FT NON TER
SO SEQUENCE 23 AA; 2485 MW; C19201F5B9BE67C8 CRC64;

Query Match 27.3%; Score 30; DB 8; Length 23;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 VNVAFLHATDLLPACDGE 19
| | | | | : : | :
| | | | | : : | :
Db 4 VGIIFTVA AFLIASTSGK 21

RESULT 8
ID Q9TJ44 PRELIMINARY; PRT; 23 AA.
AC Q9TJ44;
DT 01-MAY-2000 (TREMBLrel. 13; Created)
DT 01-MAY-2000 (TREMBLrel. 13; Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13; Last annotation update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies koreana.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBL_taxonomy=97170;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsunmura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on

```

DR	chlороplast DNA sequences."
RL	Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
KW	Chloroplast.
FT	NON_TER
SQ	SEQUENCE 23 AA; 2485 MW; C19201F5B9BE67C8 CRC64;
OY	2 VNYAFLLHATDILLPACDGE 19 : : :
DB	4 VGITFYTAFLIASTSGK 21
RESULT 9	
ID	09TUJ42 PRELIMINARY; PRT; 23 AA.
AC	09TUJ42:
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DM	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE	P700 APOPROTEIN IB (FRAGMENT).
GN	PSAB.
OS	Abies lasiocarpa.
OC	Chloroplast.
CC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
NC	Coniferopsida; Coniferales; Pinaceae; Abies.
CX	NCBI_TaxId=34340;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Suyama Y., Tsunuma Y., Yoshimaru H.;
RT	"Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RL	chloroplast DNA sequences.";
DR	Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
KW	Chloroplast.
FT	NON_TER
SQ	SEQUENCE 23 AA; 2485 MW; C19201F5B9BE67C8 CRC64;
OY	2 VNYAFLLHATDILLPACDGE 19 : : :
DB	4 VGITFYTAFLIASTSGK 21
RESULT 10	
ID	09TUJ40 PRELIMINARY; PRT; 23 AA.
AC	09TUJ40:
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DM	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE	P700 APOPROTEIN IB (FRAGMENT).
GN	PSAB.
OS	Abies mariesii.
OC	Chloroplast.
CC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
NC	Coniferopsida; Coniferales; Pinaceae; Abies.
CX	NCBI_TaxId=78263;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Suyama Y., Tsunuma Y., Yoshimaru H.;
RT	"Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RL	chloroplast DNA sequences.";
DR	Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
KW	Chloroplast.
FT	NON_TER
SQ	SEQUENCE 23 AA; 2485 MW; C19201F5B9BE67C8 CRC64;
OY	2 VNYAFLLHATDILLPACDGE 19 : : :
DB	4 VGITFYTAFLIASTSGK 21
Query Match	27.3%; Score 30; DB 8; Length 23;
Best Local Similarity	33.3%; Pred. No. 2.5e+02;
Matches 6; Conservative	4; Mismatches 8; Indels 0; Gaps 0;

KW Chloroplast. 1
FT NON_TER 1
SQ SEQUENCE 23 AA: 2485 MW: C19201F5B9BE67C8 CRC64;
Query Match 27.3%; Score 30; DB 8; Length 23;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
OY 2 VNVAFLHATDLPACDGE 19
DB 4 VGIFTYAFLIASTSGK 21
RESULT 11
O9TJ38 PRELIMINARY; PRT; 23 AA.
AC O9TJ38;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies nephrolepis.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=97171;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsumura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029679; BAA88010.1; -
KW Chloroplast.
FT NON_TER 1
SQ SEQUENCE 23 AA: 2485 MW: C19201F5B9BE67C8 CRC64;
Query Match 27.3%; Score 30; DB 8; Length 23;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
OY 2 VNVAFLHATDLPACDGE 19
DB 4 VGIFTYAFLIASTSGK 21
RESULT 12
O9TJ36 PRELIMINARY; PRT; 23 AA.
AC O9TJ36;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies sachalinensis.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=78264;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsumura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029680; BAA88012.1; -
KW Chloroplast.
FT NON_TER 1
SQ SEQUENCE 23 AA: 2485 MW: C19201F5B9BE67C8 CRC64;

Query Match 27.3%; Score 30; DB 8; Length 23;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
OY 2 VNVAFLHATDLPACDGE 19
DB 4 VGIFTYAFLIASTSGK 21
RESULT 13
O9TJ34 PRELIMINARY; PRT; 23 AA.
AC O9TJ34;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies sibirica.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=97169;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsumura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029681; BAA88014.1; -
KW Chloroplast.
FT NON_TER 1
SQ SEQUENCE 23 AA: 2485 MW: C19201F5B9BE67C8 CRC64;
Query Match 27.3%; Score 30; DB 8; Length 23;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
OY 2 VNVAFLHATDLPACDGE 19
DB 4 VGIFTYAFLIASTSGK 21
RESULT 14
O9TJ32 PRELIMINARY; PRT; 23 AA.
AC O9TJ32;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE P700 APOPROTEIN IB (FRAGMENT).
GN PSAB.
OS Abies veitchii.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=78262;
RN [1]
RP SEQUENCE FROM N.A.
RA Suyama Y., Tsumura Y., Yoshimaru H.;
RT "Molecular phylogenetic position of Japanese Abies (Pinaceae) based on
RT chloroplast DNA sequences.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029682; BAA88016.1; -
KW Chloroplast.
FT NON_TER 1
SQ SEQUENCE 23 AA: 2485 MW: C19201F5B9BE67C8 CRC64;
Query Match 27.3%; Score 30; DB 8; Length 23;

Best Local Similarity: 33.3%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 VNAYFLATDLPACDGE 19
1 | | | : : :
4 VGIITYAFLAFLISTSGK 21

RESULT 15

ID 007916 PRELIMINARY: PRT: 25 AA.
AC 007916;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)
DE HYPOTHETICAL 2.6 KDA PROTEIN.
OS Mycobacterium tuberculosis, and Mycobacterium bovis BCG.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 33892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-2296207, AND PASTEUR;
RA Supply P., Magdalena J., Himpens S., Loch C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y13628; CAA73960.1; -;
DR EMBL: Y13627; CAA73959.1; -;
DR EMBL: Y13628; CAA73959.1; -;
KW Hypothetical protein.
SQ SEQUENCE 25 AA: 2639 MW: 036E7271A33D1CF2 CRC64;

Query Match 27.3%; Score 30; DB 2; Length 25;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 PACDGE 19
1 | | | : : :
18 PACGGE 23

RESULT 16

ID 015988 PRELIMINARY: PRT: 20 AA.
AC 015988;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, last annotation update)
DE HLX1 PROTEIN (FRAGMENT).
GN HLX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93194183; PubMed=8095486;
RA Nishimura D.Y., Purchio A.F., Murray J.C.;
RT "Linkage localization of TGFB2 and the human homeobox gene HLX1 to
chromosome 1q."
RL Genomics 15:357-364(1993).
DR EMBL: S56767; AAD13883.1; -;
FT NON_TER 1 1
SQ SEQUENCE 20 AA: 2012 MW: 6BB655F09B5B5A4 CRC64;

Query Match 26.4%; Score 29; DB 4; Length 20;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 14 PACDGER 20
1 | | | : : :
7 PAADGEO 13

RESULT 17

ID 034195 PRELIMINARY: PRT: 22 AA.
AC 034195;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)
DE GROES (FRAGMENT).
GN GROES.
OS Ehrlichia risticii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Ehrlichieae; Ehrlichia.
OX NCBI_TaxID=950;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97373904; PubMed=9230387;
RA Sumner J.W., Nicholson W.L., Massung R.F.;
RT PCR amplification and comparison of nucleotide sequences from the
groEL heat shock operon of Ehrlichia species."
RL J. Clin. Microbiol. 35:2087-2092(1997).
DR EMBL: U96732; AAB65632.1; -;
DR INTERPRO: IPR001476; -;
DR PFM: PFM00166; cpn10; 1.
FT NON_TER 1 1
SQ SEQUENCE 22 AA: 2504 MW: 21D959D75191F620 CRC64;

Query Match 26.4%; Score 29; DB 2; Length 22;
Best Local Similarity 27.8%; Pred. No. 3.6e+02;
Matches 5; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 VNAYFLATDLPACDGE 19
1 | | | : : :
5 VEYVWKEDIIAKKSG 22

RESULT 18

ID 085713 PRELIMINARY: PRT: 15 AA.
AC 085713;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, last annotation update)
DE 29 (RSV-29) SRC (FRAGMENT).
OS Rous sarcoma virus.
OC Viruses; Retrovirus; Retroviridae; Avian type C retroviruses.
OX NCBI_TaxID=11886;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85265036; PubMed=2991593;
RA Dutta A., Wang L.H., Hanafusa T., Hanafusa H.;
RT "Partial nucleotide sequence of Rous sarcoma virus-29 provides
evidence that the original Rous sarcoma virus was replication
defective."
RL J. Virol. 55:728-735(1985).
DR EMBL: M1117; AAA42556.1; -;
FT NON_TER 1 1
SQ SEQUENCE 15 AA: 1597 MW: DB53E0F31273C5C1 CRC64;

Query Match 25.5%; Score 28; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LIPAC 16
1 | | | : : :
5 LIPAC 9

RESULT 19

Q9S8J0

ID 095830 PRELIMINARY; PRT: 21 AA.
 AC 095830;
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)
 DE 27 KDA RIBOSOME-INACTIVATING PROTEIN (FRAGMENT).
 OS Bryonia dioica (Red bryon).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Cucurbitales; Cucurbitaceae; Bryonia.
 NCBI_TaxID=3652;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95151812; PubMed=7849072;
 RA Siegel C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
 RA Marquardt H.;
 RT "Characterization of ribosome-inactivating proteins isolated from
 RT Bryonia dioica and their utility as carcinoma-reactive
 RT immunocjugates.";
 RL Bioconjug. Chem. 5:423-429(1994).
 SO SEQUENCE 21 AA; 2302 MW; 54ED5E00AD12787B CRC64;

Query Match 25.5%; Score 28; DB 10; Length 21;
 Best Local Similarity 40.0%; Pred. No. 5e+02; Mismatches 2; Indels 0; Gaps 0;
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 DYNATFLHAT 10
 ID 09TRC6 PRELIMINARY; PRT: 23 AA.
 AC 09TRC6;
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)
 DE CRYOCHROME P450 DPB-1 ISOFORM (FRAGMENT).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94198820; PubMed=8185738;
 RA Shiraga T., Iwasaki K., Nozaki K., Yamura T., Yamazoe Y., Kato R.,
 RA Takahaka A.;
 RT "Isolation and characterization of four cytochrome P450 isozymes from
 RT untreated and phenobarbital-treated beagle dogs.";
 RL Biol. Pharm. Bull. 17:22-28(1994).
 SO SEQUENCE 23 AA; 2638 MW; AF86888FE8029F8 CRC64;

Query Match 25.5%; Score 28; DB 6; Length 23;
 Best Local Similarity 55.6%; Pred. No. 5.5e+02; Mismatches 2; Indels 0; Gaps 0;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 DLPACDGE 19
 ID 094377 PRELIMINARY; PRT: 24 AA.
 AC 094377;
 DT 01-FEB-1997 (TRENBLREL. 02, Created)
 DT 01-FEB-1997 (TRENBLREL. 02, Last sequence update)
 DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
 DE PPI-LIKE SER/THR PROTEIN PHOSPHATASE (FRAGMENT).
 OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zeke T., Gergely P., Dombradi V.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 277740; CAB01299.1; -.
 DR HSSP; P08129; 150M.
 DR INTERPRO; IPR000934; -.
 DR PFAM; PF00149; Stposphatase; 1.
 FT NON_TER 1 1
 FT NON_TER 24 24
 SO SEQUENCE 24 AA; 2816 MW; 2BF73DBB5F20359 CRC64;

Query Match 25.5%; Score 28; DB 5; Length 24;
 Best Local Similarity 71.4%; Pred. No. 5.7e+02; Mismatches 1; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DYNATFL 7
 ID 09T055 PRELIMINARY; PRT: 21 AA.
 AC 09T055;
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
 DE UDP-GLUCURONOSYLTRANSFERASE 1A1 (FRAGMENT).
 GN UGT1A1.
 OS Cebus apella (Brown-capped capuchin).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Cebus.
 NCBI_TaxID=9515;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hall D., Ybaze G., Destro-Bisol G., Petal-Erler M.L., Di Rienzo A.;
 RT "Variability at the uridine diphosphate glucuronosyltransferase 1A1
 RT promoter in human populations and primates.";
 RL Pharmacogenetics 0:0-0(1999).
 DR EMBL; AF135470; AF09181.1; -.
 KM Transferrase.
 FT NON_TER 21 21
 FT NON_TER 21 21
 SO SEQUENCE 21 AA; 2000 MW; 9887EF96C931C820 CRC64;

Query Match 24.5%; Score 27; DB 6; Length 21;
 Best Local Similarity 66.7%; Pred. No. 7.3e+02; Mismatches 1; Indels 0; Gaps 0;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 LPACDG 18
 ID 030216 PRELIMINARY; PRT: 18 AA.
 AC 030216;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)
 DE HUMAN LECTOCYTE ANTIGEN ALPHA CHAIN (FRAGMENT).
 GN HLA DOA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-95064785; PubMed-7974465;
 RA Fogdell A., Olerup O.;
 RT "The DAI*0104 allele is carried by DRB1*1001- and DRB1*1401-positive
 RT haplotypes in Caucasians, Africans and Orientals."
 RT Tissue Antigens 44:19-24(1994).
 DR EMBL: S75685; AAB32621.1; -
 DR INTERPRO: IPR001003; -
 DR PFM: PF00993; MHC-II_alpha: 1.
 FT NON_TER
 SQ SEQUENCE 18 AA; 1708 MW; 097492525C2252FC CRC64;

Query Match 23.6%; Score 26; DB 7; Length 18;
 Best Local Similarity 38.5%; Pred. No. 9.1e+02;
 Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 LHATDLPACDGE 19
 | | : | | |
 DB 3 LALTTMSPCGGE 15

RESULT 24
 OSUMM9 PRELIMINARY; PRT; 19 AA.
 AC Q9UMM9:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE ARGININE VASOPRESSIN-NEUROPHYSIN II (FRAGMENT).
 GN AVP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91123474; PubMed-1840604;
 RA Ito M., Mori Y., Oiso Y., Saito H.;
 RT "A single base substitution in the coding region for neurophysin II
 RT associated with familial central diabetes insipidus."
 RL J. Clin. Invest. 87:725-728(1991).
 DR EMBL: M63733; AAA69844.1; -
 FT NON_TER
 SQ SEQUENCE 19 AA; 1985 MW; 5FF5FCD7BD990451 CRC64;

Query Match 23.6%; Score 26; DB 4; Length 19;
 Best Local Similarity 80.0%; Pred. No. 9.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 LFPAC 16
 : | | | |
 DB 5 MLPAC 9

RESULT 25
 OSUMM9 PRELIMINARY; PRT; 20 AA.
 AC Q9QVG0:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CLEAVED PROLACTIN-1, CLPRL-1-FRAGMENT A.
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID-10118;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-92143803; PubMed-1736889;
 RA Andries M., Tillemans D., Deneef C.;
 RT "Isolation of cleaved prolactin variants that stimulate DNA synthesis

RT in specific cell types in rat pituitary cell aggregates in culture."
 RL Biochem. J. 281:393-400(1992).
 SQ SEQUENCE 20 AA; 2146 MW; 89CA019A7668CBBA CRC64;

Query Match 23.6%; Score 26; DB 11; Length 20;
 Best Local Similarity 66.7%; Pred. No. 1e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 LPACDG 18
 | | | |
 DB 1 LPVCSC 6

Search completed: February 5, 2001, 10:55:20
 Job time: 909 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:37:07; Search time 58.67 Seconds
(without alignments)
23.147 Million cell updates/sec

Title: us-08-981-824-1

Sequence: 1 DVNRYAFHARDLPACDGER 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 5001

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 66: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	28.2	25	2	S10850
2	29	26.4	16	2	PH1640
3	28	25.5	20	2	S15861
4	28	25.5	24	2	T42014
5	26	23.6	18	2	I59649
6	26	23.6	19	2	H27480
7	26	23.6	24	2	T42257
8	26	23.6	24	2	T42441
9	25	22.7	18	2	S58277
10	25	22.7	20	2	S65884
11	25	22.7	23	2	B39855
12	25	22.7	23	2	S38991
13	25	22.7	24	2	T42258
14	25	22.7	25	2	S13996
15	25	22.7	25	2	S13591
16	24	21.8	11	2	S61797
17	24	21.8	16	2	S68730
18	24	21.8	18	1	QHRC2
19	24	21.8	18	2	A60103
20	24	21.8	18	2	I51427
21	24	21.8	19	2	S63489
22	24	21.8	20	2	S63490
23	24	21.8	23	2	A39855
24	24	21.8	24	2	T42259
25	23	20.9	10	2	A32543
26	23	20.9	10	2	A56533
27	23	20.9	17	2	A54534
28	23	20.9	19	2	I40063
29	23	20.9	22	2	I51830

30	23	20.9	24	2	T42256
31	23	20.9	24	2	I54750
32	22	20.0	12	2	S16204
33	22	20.0	14	2	PL0040
34	22	20.0	15	2	S30608
35	22	20.0	15	2	S55312
36	22	20.0	17	2	J02310
37	22	20.0	17	2	J02320
38	22	20.0	19	1	EW5MAN
39	22	20.0	23	2	D34047
40	22	20.0	23	2	B42382
41	22	20.0	24	2	T42378
42	22	20.0	25	2	S20257
43	21	19.1	11	2	B49164
44	21	19.1	12	2	S01122
45	21	19.1	13	2	S01119

ALIGNMENTS

RESULT 1
S10850
alpha-amylase inhibitor - durum wheat (fragment)
N: Alternate names: glutenin low molecular weight chain
C: Species: Triticum durum (durum wheat)
C: Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
C: Accession: S10850
R: Kobrehel, K.; Alary, R.
J. Sci. Food Agric. 48, 441-452, 1989
A: Title: Isolation and partial characterization of two low molecular weight durum whe
A: Reference number: S10849
A: Accession: S10850
A: Molecule type: protein
A: Residues: 1-25 <KOB>
C: Superfamily: wheat alpha-amylase inhibitor
C: Keywords: alpha-amylase inhibitor

Query Match 28.2%; Score 31; DB 2; Length 25;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 TDLIPAC 16
DB 12 TBLPAC 18

RESULT 2
PH1640
Ig H chain V-D-J region (clone B-lees 235) - mouse (fragment)
C: Species: Mus musculus (house mouse)
C: Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C: Accession: PH1640
R: Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A: Title: Molecular characterization of transgene-induced immunodeficiency in B-lees m
A: Reference number: PH1580; MUID: 93301609
A: Accession: PH1640
A: Molecule type: DNA
A: Residues: 1-16 <LEV>
A: Experimental source: bone marrow pre-B lymphocyte
C: Keywords: immunoglobulin

Query Match 26.4%; Score 29; DB 2; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 5 AFLHATLPPACDG 18
DB 2 ARLRATVTPICVG 15

```

RESULT 3
S15861
estrogen receptor - p1g (fragments)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 30-May-1997
C:Accession: S15861
R:Inole, H.H.; Jungblut, P.W.; Jakob, F.
Biochem. J. 276, 709-714, 1991
A:Title: The proton-driven dissociation of oestradiol-receptor dimers as a preparative t
9.
A:Reference number: S15861; MUID:91291128
A:Accession: S15861
A:Molecule type: protein
A:Residues: 1-20 <Bio>
C:Keywords: steroid hormone receptor

Query Match      25.5%; Score 28; DB 2; Length 20;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNVAFLHA 9
Db 6 INMAKLHA 13

RESULT 4
T42014
Ser/Thr protein phosphatase, PP1-like - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: T42014
R:Zeke, T.; Gergely, P.; Dombradi, V.
Submitted to the EMBL Data Library, July 1996
A:Description: The catalytic subunits of Ser/Thr protein phosphatases from Caenorhabdit
A:Reference number: 222025
A:Accession: T42014
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-24 <Zek>
A:Cross-references: EMBL:277740; PIDN:CAB01299.1
A:Experimental source: strain Bristol; clone C-1
C:Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprotei

Query Match      25.5%; Score 28; DB 2; Length 24;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DVNYAFL 7
Db 17 DVNYFL 23

RESULT 5
I59649
human leukocyte antigen alpha chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I59649
R:Fogdell, A.; Olerup, O.
Tissue Antigens 44, 19-24, 1994
A:Title: The DOA1*0104 allele is carried by DRB1*1001- and DRB1*1401-positive haplotypes
A:Reference number: I59649; MUID:35064785
A:Accession: I59649
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-18 <RES>
A:Cross-references: GB:S75665; MID:9913771; PIDN:AA32621.1; PID:9913772
C:Genetics:
A:Gene: GDB:HLA-DOA1
A:Cross-references: GDB:120638; OMIM:146880

```

```

A:Map position: 6p21.3-6p21.3

Query Match      23.6%; Score 26; DB 2; Length 18;
Best Local Similarity 38.5%; Pred. No. 6e+02;
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 LHA7DLPAQDGE 19
Db 3 LALTRMSPGCGE 15

RESULT 6
H27480
hydrogenase (EC 1.18.99.1) large chain - Desulfovibrio baculatus (strain Norway 4) (f
N:Alternate names: hydrogenylase; Nitase hydrogenase
C:Species: Desulfovibrio baculatus
C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 30-Sep-1993
C:Accession: H27480
R:Prickrill, B.C.; He, S.H.; Li, C.; Menon, N.; Chol, E.S.; Przybyla, A.E.; DerVartan
Biochem. Biophys. Res. Commun. 149, 369-377, 1987
A:Title: Identification of three classes of hydrogenase in the genus, Desulfovibrio.
A:Reference number: A27480; MUID:88106446
A:Accession: H27480
A:Molecule type: protein
A:Residues: 1-19 <PR1>
C:Keywords: oxidoreductase

Query Match      23.6%; Score 26; DB 2; Length 19;
Best Local Similarity 57.1%; Pred. No. 6.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 14 PACDGER 20
Db 6 PAADGKR 12

RESULT 7
T42257
phosphoprotein phosphatase (EC 3.1.3.16) - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: T42257
R:Zeke, T.; Gergely, P.; Dombradi, V.
Submitted to the EMBL Data Library, July 1996
A:Description: The catalytic subunits of Ser/Thr protein phosphatases from Caenorhabd
A:Reference number: 222131
A:Accession: T42257
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-24 <Zek>
A:Cross-references: EMBL:277735; PIDN:CAB01294.1
C:Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprot
C:Keywords: phosphoric monoester hydrolase

Query Match      23.6%; Score 26; DB 2; Length 24;
Best Local Similarity 83.3%; Pred. No. 8.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNVAFL 7
Db 18 VNVAFL 23

RESULT 8
T42441
protein phosphatase - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: T42441
R:Zeke, T.; Gergely, P.; Dombradi, V.

```

submitted to the EMBL Data Library, July 1996

A:Description: The catalytic subunits of Ser/Thr protein phosphatases from *Caenorhabditis*

A:Reference number: 222131

C:Date: 14-Feb-1997

C:Accession: S55884

R:Kasperaitis, M.A.M.; Voorma, H.O.; Thomas, A.A.M.

FEBS Lett. 365, 47-50, 1995

A:Title: The amino acid sequence of eukaryotic translation initiation factor 1 and its

A:Reference number: S55884; MUID:95293122

A:Accession: S55884

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-24 <2EK>

A:Cross-references: EMBL:277729; PIDN:CA01288.1

C:Superfamily: phosphoprotein phosphatase; phosphodiesterase core homology; phosphoprotein

Query Match 23.6%; Score 26; DB 2; Length 24;

Best Local Similarity 83.3%; Pred. No. 8.1e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 VNYAFL 7

Db 18 VNYLFL 23

RESULT 9

S58277

Insulin-like growth factor receptor type II - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999

C:Accession: S58277

R:Smirka, O.W.; Stoger, R.; Kurzbauer, R.; Fae, I.; Fischer, G.F.; Barlow, D.P.

submitted to the EMBL Data Library, January 1995

A:Description: Conservation of a methylation imprint and a putative imprinting box at th

A:Reference number: S58277

A:Accession: S58277

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-18 <SKR>

A:Cross-references: EMBL:X83702; NID:9929644; PIDN:CA58675.1; PID:9929645

C:Keywords: growth factor receptor

Query Match 22.7%; Score 25; DB 2; Length 18;

Best Local Similarity 57.1%; Pred. No. 8.8e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 13 LPACDGE 19

Db 11 LPSCVCK 17

RESULT 10

S55884

translation initiation factor eIF-1 - rabbit (fragments)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 07-May-1999

C:Accession: S55884

R:Kasperaitis, M.A.M.; Voorma, H.O.; Thomas, A.A.M.

FEBS Lett. 365, 47-50, 1995

A:Title: The amino acid sequence of eukaryotic translation initiation factor 1 and its

A:Reference number: S55884; MUID:95293122

A:Accession: S55884

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-9;10-16;17-20 <KAS>

Query Match 22.7%; Score 25; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 9.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 DLPPA 15

Db 2 DLPPA 6

RESULT 11

B39855

paralytic peptide II - tobacco hornworm

C:Species: Manduca sexta (tobacco hornworm)

C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 03-Feb-1994

C:Accession: B39855

R:Skinner, W.S.; Dennis, P.A.; Li, J.P.; Summerfelt, R.M.; Carney, R.L.; Qvistad, G.B

J. Biol. Chem. 266, 12873-12877, 1991

A:Title: Isolation and identification of paralytic peptides from hemolymph of the lep

A:Reference number: A39855; MUID:91302298

A:Accession: B39855

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-23 <SKI>

C:Superfamily: paralytic peptide I

Query Match 22.7%; Score 25; DB 2; Length 23;

Best Local Similarity 43.8%; Pred. No. 1.1e+03;

Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 3 NYAFLHATDLPCDQ 18

Db 2 NFAGCATGFLRTADG 17

RESULT 12

S38991

glycine reductase (EC 1.4.99.-) sulphydryl protein C - *Eubacterium acidaminophilum* (f

N:Alternate names: glycyl reductase 57K protein

C:Species: *Eubacterium acidaminophilum*

C:Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 15-Aug-1997

C:Accession: S38991; S21250

R:Luubbers, M.; Andreesen, J.R.

Eur. J. Biochem. 217, 791-798, 1993

A:Title: Components of glycine reductase from *Eubacterium acidaminophilum*. Cloning, s

A:Reference number: S38988; MUID:94039119

A:Accession: S38991

A:Molecule type: DNA

A:Residues: 1-23 <LDE>

A:Cross-references: GB:L04500

R:Schraeder, T.; Andreesen, J.R.

Eur. J. Biochem. 206, 79-85, 1992

A:Title: Purification and characterization of protein P(c), a component of glycine re

A:Reference number: S21222; MUID:92267030

A:Accession: S21250

A:Molecule type: protein

A:Residues: 1-23 <SCH>

C:Genetics:

A:Gene: grdc

C:Function:

A:Description: glycine reductase complex catalyzes the reductive deamination of glyci

C:Keywords: ATP; oxidoreductase

Query Match 22.7%; Score 25; DB 2; Length 23;

Best Local Similarity 30.0%; Pred. No. 1.1e+03;

Matches 3; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 4 YAFLNATDGL 13

Db 11 YVLVHTPDWI 20

RESULT 13

T42258

Ser/Thr protein phosphatase homolog - *Caenorhabditis elegans* (fragment)

C:Species: *Caenorhabditis elegans*

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000

C:Accession: T42258

R:Zeke, T.; Gergely, P.; Dombadi, V.

submitted to the EMBL Data Library, July 1996

A:Description: The catalytic subunits of Ser/Thr protein phosphatases from *Caenorhabd*

A:Reference number: 222131

A:Accession: T42258
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-24 <ZK>
 A:Cross-references: EMBL:Z77734; PIDN:CAB01293.1
 C:Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprotein

Query Match 22.7%; Score 25; DB 2; Length 24;
 Best Local Similarity 36.4%; Pred. No. 1.2e+03;
 Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 7 LHATDLPACD 17
 | | | | |
 Db 3 IHLRLMTCD 13

RESULT 14
 S13996
 hypothetical protein - phage TW19 (fragment)
 C:Species: phage TW19
 C:Date: 18-Feb-1994 #sequence_revision 24-Apr-1998 #text_change 24-Apr-1998
 C:Accession: S13996
 R:Inokuchi, Y.; Hirashima, A.; Matanabe, I.
 J. Mol. Biol. 158, 711-730, 1982
 A:Title: Comparison of the nucleotide sequences at the 3'-terminal region of RNAs from F
 A:Reference number: S07250; MUID:83010313
 A:Accession: S13996
 A:Status: translation not shown
 A:Molecule type: genomic RNA
 A:Residues: 1-25 <INO>
 A:Cross-references: EMBL:J05520; NID:9216180

Query Match 22.7%; Score 25; DB 2; Length 25;
 Best Local Similarity 60.0%; Pred. No. 1.2e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 14 PACDG 18
 | | | | |
 Db 12 PSCG 16

RESULT 15
 S11351
 spore storage protein light chain isoform M2b - ostrich fern (fragment)
 N:Alternate names: matuccin
 C:Species: Matuccia struthiopteris (ostrich fern)
 C:Date: 19-Mar-1997 #sequence_revision 21-Nov-1998 #text_change 01-Feb-1999
 C:Accession: S11351
 R:Roedin, J.; Rask, L.
 Eur. J. Biochem. 192, 101-107, 1990
 A:Title: Characterization of matuccin, the 2.2S storage protein of the ostrich fern.
 A:Reference number: S11350; MUID:90382429
 A:Accession: S11351
 A:Molecule type: protein
 A:Residues: 1-25 <ROE>
 C:Keywords: heterodimer; seed

Query Match 22.7%; Score 25; DB 2; Length 25;
 Best Local Similarity 62.5%; Pred. No. 1.2e+03;
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 LHATDLP 14
 | | | | |
 Db 12 LHQCDLP 19

RESULT 16
 S61797
 T-cell-specific transcription factor 1 splice form F - human (fragment)
 N:Alternate names: transcription factor TCF-1F

C:Species: Homo sapiens (man)
 C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 24-Jul-1998
 C:Accession: S61797; S61879
 R:Mayer, K.; Wolff, E.; Clevers, H.; Ballhausen, W.G.
 Biochim. Biophys. Acta 1263, 169-172, 1995
 A:Title: The human high mobility group (HMG)-box transcription factor TCF-1: novel is
 A:Reference number: S61796; MUID:95367594
 A:Accession: S61797
 A:Molecule type: mRNA
 A:Residues: 1-11 <MAV>
 A:Cross-references: EMBL:Z47363
 A:Note: DNA was also sequenced
 C:Keywords: alternative splicing; DNA binding; transcription factor

Query Match 21.8%; Score 24; DB 2; Length 11;
 Best Local Similarity 66.7%; Pred. No. 7.8e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 DLTPAC 16
 | | | | |
 Db 1 DGPAC 6

RESULT 17
 S68730
 bleomycin-binding protein - Streptomyces verticillius (fragment)
 C:Species: Streptomyces verticillius
 C:Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 17-Mar-1999
 C:Accession: S68730
 R:Sugiyama, M.; Kumagai, T.; Matsuo, H.; Bhuiyan, M.Z.A.; Ueda, K.; Mochizuki, H.; Na
 FEBS Lett. 362, 80-84, 1995
 A:Title: Overproduction of the bleomycin-binding proteins from bleomycin-producing St
 al. characterization.
 A:Reference number: S68730; MUID:95212588
 A:Accession: S68730
 A:Molecule type: protein
 A:Residues: 1-16 <SUG>
 A:Experimental source: ATCC 15003
 C:Keywords: antibiotic resistance

Query Match 21.8%; Score 24; DB 2; Length 16;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 FLHATDLPACD 17
 | | | | |
 Db 4 FLGAVPVLTAVD 15

RESULT 18
 QHEC2
 heat-stable enterotoxin ST-2 - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 06-Jul-1982 #sequence_revision 06-Jul-1982 #text_change 31-Dec-1996
 C:Accession: A01823
 R:Chan, S.K.; Giannelis, R.A.
 J. Biol. Chem. 256, 7744-7746, 1981
 A:Title: Amino acid sequence of heat-stable enterotoxin produced by Escherichia coli
 A:Reference number: A01823; MUID:81264141
 A:Accession: A01823
 A:Molecule type: protein
 A:Residues: 1-18 <CHA>
 A:Experimental source: strain 18D, serotype 0.42:H6:H37
 C:Comment: This enterotoxin is one of several, of differing molecular sizes, produced
 idues of the heat-stable enterotoxin ST-1.
 C:Superfamily: heat-stable enterotoxin ST
 C:Keywords: enterotoxin; heat-stable protein
 F:1-18/Product: heat-stable enterotoxin ST-2 #status experimental <MAT>
 F:3-10,6-14,9-17/Disulfide bonds: #status predicted

Query Match 21.8%; Score 24; DB 1; Length 18;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 PACDG 18
||| |
DB 12 PACAG 16

RESULT 19
A:Accession: A60103
heat-stable enterotoxin ST-Ia - Citrobacter freundii
C:Species: Citrobacter freundii
C:Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 18-Jun-1993
C:Accession: A60103
R:Guarino, A.; Giannella, R.; Thompson, M.R.
Infect. Immun. 57, 649-652, 1989
A:Title: Citrobacter freundii produces an 18-amino-acid heat-stable enterotoxin identical
A:Reference number: A60103; MUID:89108617
A:Accession: A60103
A:Molecule type: protein
A:Residues: 1-18 <CUA>
C:Superfamily: heat-stable enterotoxin ST

Query Match 21.8%; Score 24; DB 2; Length 18;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 PACDG 18
||| |
DB 12 PACAG 16

RESULT 20
A:Accession: I51427
hemoglobin alpha chain - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 19-May-2000
C:Accession: I51427
R:Patent, R.K.; Elkington, J.A.; Kay, R.M.; Williams, J.G.
Cell 21, 565-573, 1980
A:Title: Internal organization of the major adult alpha- and beta-globin genes of X.laevis
A:Reference number: I51427; MUID:81001900
A:Accession: I51427
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-18 <PAT>
A:Cross-references: GB:J00975; NID:g214198; PID:g214199
C:Superfamily: globin; globin homology
C:Keywords: blood; oxygen carrier

Query Match 21.8%; Score 24; DB 2; Length 18;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 LHATDL 12
||| |
DB 6 LHAIDL 11

RESULT 21
S63489
disulfite reductase alpha chain, membrane-bound - Desulfovibrio desulfuricans
C:Species: Desulfovibrio desulfuricans
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S63489
R:Seuwer, J.; Arendsen, A.F.; Hagen, W.R.; Kroneck, P.M.H.
Eur. J. Biochem. 233, 873-879, 1995
A:Title: Molecular properties of the disulfite reductase from Desulfovibrio
A:Reference number: S63489; MUID:96085152
A:Accession: S63489

A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-19 <STE>

Query Match 21.8%; Score 24; DB 2; Length 19;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 HATDL 13
||| |
DB 3 HATPL 8

RESULT 22
S63490
disulfite reductase alpha chain, soluble - Desulfovibrio desulfuricans (f
C:Species: Desulfovibrio desulfuricans
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S63490
R:Seuwer, J.; Arendsen, A.F.; Hagen, W.R.; Kroneck, P.M.H.
Eur. J. Biochem. 233, 873-879, 1995
A:Title: Molecular properties of the disulfite reductase from Desulfovibrio
A:Reference number: S63489; MUID:96085152
A:Accession: S63490
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <STE>

Query Match 21.8%; Score 24; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 HATDL 13
||| |
DB 3 HATPL 8

RESULT 23
A39855
paralytic peptide I - tobacco hornworm
C:Species: Manduca sexta (tobacco hornworm)
C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 03-Feb-1994
C:Accession: A39855
R:Skinner, W.S.; Dennis, P.A.; Li, J.P.; Summerfelt, R.M.; Carney, R.L.; Quistad, G.B.
J. Biol. Chem. 266, 12873-12877, 1991
A:Title: Isolation and identification of paralytic peptides from hemolymph of the lep
A:Reference number: A39855; MUID:91302298
A:Accession: A39855
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-23 <SKI>
C:Superfamily: paralytic peptide I

Query Match 21.8%; Score 24; DB 2; Length 23;
Best Local Similarity 43.8%; Pred. No. 1.7e+03;
Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 NVAFLHATDLIPACDG 18
||| |
DB 2 NFAGCATGYLRADG 17

RESULT 24
T42259
phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain - Caenorhabditis elegans (f
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: T42259
R:Zeke, T.; Gergely, P.; Dombradi, V.
submitted to the EMBL Data Library, July 1996

A:Description: The catalytic subunits of Ser/Thr protein phosphatases from Caenorhabditis
A:Reference number: 222131
A:Accession: T42259
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-24 <ZEK>
A:Cross-references: EMBL:Z77733; PIDN:CAB01292.1
C:Genetics:
A:Note: PPI
C:Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprotein
C:Keywords: phosphoric monoester hydrolase

Query Match 21.8%; Score 24; DB 2; Length 24;
Best Local Similarity 57.1%; Pred. NO. 1.7e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 DVNYAFL 7
: ||||
Db 17 EANYLFL 23

RESULT 25
A32543
cardioexcitatory neuropeptide - desert locust
C:Species: Schistocerca gregaria (desert locust)
C:Date: 20-Dec-1989 #sequence_revision 20-Dec-1989 #text_change 18-Aug-2000
C:Accession: A32543
R:Robb, S.; Packman, L.C.; Evans, P.D.
Biochem. Biophys. Res. Commun. 160, 850-856, 1989
A:Title: Isolation, primary structure and bioactivity of Schistoflrf-amide, a FMRF-amide
A:Reference number: A32543; MUID:89246543
A:Accession: A32543
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <ROB>
C:Superfamily: unassigned animal peptides
C:Keywords: amidated carboxyl end; neuropeptide
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 20.9%; Score 23; DB 2; Length 10;
Best Local Similarity 57.1%; Pred. NO. 1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DVNYAFL 7
||:|
Db 2 DVDHYFL 8

Search completed: February 5, 2001, 10:49:26
Job time: 739 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:37:06 ; Search time 68.99 Seconds
(without alignments)
9.913 Million cell updates/sec

Title: US-08-981-824-1

Perfect score: 110
Sequence: 1 DVNVAFLHATDLPACDGER 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_36:*

1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	20	18	W18842
2	110	100.0	20	18	W01793
3	73	66.4	14	16	R76649
4	73	66.4	14	18	W18857
5	73	66.4	20	16	R72267
6	73	66.4	20	21	Y59569
7	58.2	58.2	20	16	R72266
8	58.2	58.2	20	21	Y59540
9	64	58.2	20	21	Y59568
10	38	34.5	19	19	W79176
11	34	30.9	15	14	R43901
12	34	30.9	15	19	W58542

13	34	30.9	15	20	W87621
14	34	30.9	15	20	W72822
15	34	30.9	20	19	W76928
16	32	29.1	9	21	Y82873
17	32	29.1	15	19	W85260
18	32	29.1	20	9	P83057
19	31	28.2	17	13	R23839
20	31	28.2	19	13	R23840
21	31	28.2	21	20	W96277
22	30	27.3	7	18	W30416
23	30	27.3	15	17	R89429
24	30	27.3	15	19	W85278
25	30	27.3	20	16	R36614
26	30	27.3	20	16	R84360
27	30	27.3	20	16	R84361
28	29	26.4	10	19	W68470
29	29	26.4	10	20	Y25488
30	29	26.4	10	20	Y07180
31	29	26.4	15	21	Y67116
32	29	26.4	15	21	Y67117
33	29	26.4	15	21	Y54753
34	29	26.4	15	21	Y54754
35	29	26.4	19	11	R04826
36	29	26.4	22	9	P82431
37	29	26.4	22	20	Y36422
38	29	26.4	22	20	Y27745
39	28	25.5	12	21	Y83094
40	28	25.5	19	16	R97609
41	28	25.5	21	20	Y04383
42	28	25.5	21	20	Y04385
43	28	25.5	22	6	P50362
44	28	25.5	22	9	P82432
45	28	25.5	22	18	W35329

ALIGNMENTS

RESULT	1
ID	W18842
XX	W18842 standard; peptide: 20 AA.
AC	W18842;
DT	05-JAN-1998 (first entry)
XX	
DE	65 kD glutamic acid decarboxylase peptide fragment I.
XX	
KW	GAD, 65 kD; human; glutamic acid decarboxylase; autoreactive; diabetes;
KW	Insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
KW	predisposition; autoimmune; tumour; rheumatoid arthritis;
KW	multiple sclerosis.
OS	Synthetic.
XX	
PN	DE19526561-A1.
XX	
PD	23-JAN-1997.
XX	
PF	20-JUL-1995; 95DE-1026561.
XX	
PR	20-JUL-1995; 95R-1026561.
XX	
PA	(BOEF) BOEHRINGER MANNHEIM GMBH.
XX	
PI	Donté F, Endl J, Ganz M, Jung G, Kientsch-Engel R;
XX	Pozzilli P, Stahl P;
DR	WPI; 1997-088254/09.
PT	Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
XX	Involving intradermal admin. of auto-reactive substances

Epitope of HIV-1 g
HIV-1 gp120 monocl
Fusion immunoglobu
Teratocarcinoma-de
Helper T-cell pep
Pep-14 comprising
Peptide from the E
B-cell epitope of
HRE-I affinity pep
KIR-conjugated pep
Helper T-cell pep
HCV peptide NP-17.
Hepatitis C virus
Hepatitis C virus
Amino acids 250-25
Insulin-like growth
HIV binding peptide
Human protease pep
Human protease pep
Human subtilisin f
Junction of STA-re
Peptide unit of pr
Fragment of human
Human secreted pro
Cull peptide anti
Human thyroglobuli
Partial pullanase
Partial pullanase
Heparin derived cy
Pseudomonas HSP60

```

PS Claim 11; Page 9; 12pp; German.
CC W18842-70 are peptide fragments of the 65 kD human glutamic acid
CC decarboxylase (GAD). The fragments are autoreactive substances used for
CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
CC determined by using a claimed method for diagnosis of cell-mediated
CC diseases or a predisposition to cell-mediated diseases, which is effected
CC by administering an autoreactive substance intradermally and establishing
CC the diagnosis on the basis of the occurrence or lack of a positive
CC reaction at the site of administration. The method is used for diagnosis
CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
CC
XX Sequence 20 AA:
SQ
QY 1 DVNYAFHLNATDLLPACDGER 20
   |||||
Db 1 dvnyafhlhatalldpacdger 20
   |||||

RESULT 2
DE W01793 standard; peptide; 20 AA.
W01793
XX W01793:
XX
XX 15-OCT-1997 (first entry)
XX
DE Human 65 kD glutamine decarboxylase peptide.
XX
XX Human; glutamine decarboxylase; GAD; diagnosis; predisposition;
XX tumour; immunological; disease; autoimmune; diabetes; diabetes; reagent;
XX determination; T cell; subpopulation; medicament; treatment;
XX prevention; production; antigen; immunogen; tolerogen; isolation;
XX reinjection; inactivation.
XX
XX Homo sapiens.
XX
XX DE19525784-A1.
XX
XX 16-JAN-1997.
XX
XX 14-JUL-1995; 95DE-1025784.
XX
XX 14-JUL-1995; 95DE-1025784.
XX
XX (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX
XX Albert W, Boltard C, Endl J, Jung G, Schendel D;
XX Stahl P, Van Endert P;
XX
XX WPI: 1997-078452/08.
XX
XX Glutamine decarboxylase peptide(s) - for diagnosis and therapy of
XX diabetes, etc.
XX
XX Claim 1; Page 12; 15pp; German.
XX
CC The present peptide is a fragment of the human 65 kD glutamine
CC decarboxylase (GAD), which can be used to diagnose, or diagnose a
CC predisposition to, a tumour or immunological disease, preferably an
CC autoimmune disease, especially diabetes. It can also be used as a
CC reagent to determine specific T cell subpopulations, in medicaments
CC to treat or prevent immunological diseases, preferably autoimmune
CC diseases, especially diabetes, to produce antigens, especially
CC immunogens or tolerogens and to isolate specific T cell
CC subpopulations, which can be used to produce antigens or for
CC reinjection, optionally after inactivation.

```

SQ	Sequence	20 AA:
OY	Query Match	100.0%; Score 110; DB 18; Length 20;
	Best Local Similarity	100.0%; Pred. No. 4.7e-12;
Matches	20; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1 DVNYAFILHATDILLPACDGER 20 	
DB	1 dvnyafilmatdillpacdgger 20	
RESULT	3	
ID	R76649	
XX	R76649 standard; peptide; 14 AA.	
XX	R76649;	
DT	05-MAR-1996 (first entry)	
DE	Peptide derived from human glutamic acid decarboxylase 8.	
KW	diabetes; T-cell subpopulation; detection; antigen production; diagnosis; autoimmune disease.	
OS	Homo sapiens.	
PX	DE4418091-AI.	
PD	27-JUL-1995.	
PE	24-MAY-1994; 94DE-4418091.	
PR	04-FEB-1994; 94DE-4403522.	
PR	20-JAN-1994; 94DE-4401629.	
PA	(ENDL/) ENDL J. (BOEF) BOEHRINGER MANNHEIM GMBH.	
PI	Albert W, Dornmair K, Endl J, Jung G, Meinel E; Stahl P, Schendel D;	
DR	WPL, 1995-264505/35.	
PT	Antigen-specific activated T-lymphocytes and their detection - by interaction with inventive peptide(s) or peptide-MHC complexes; useful in diagnosis of e.g. diabetes and auto-immune diseases	
PS	Claim 1; Fig 2; 21pp; German.	
CC	R76642-62 are derived from human glutamic acid decarboxylase and specifically react with T-cell sub-populations isolated from recently diagnosed Type-I diabetics. Pharmaceutical compns. contg. these peptides and those shown in R77571-72, are useful for the diagnosis of a disease or predispositions of immune system diseases, tumours, and autoimmune diseases, including diabetes. The peptides are able to detect specific T-cell subpopulations that are then used for antigen prodn., e.g. by reinjection.	
SO	Sequence 14 AA:	
Query Match	66.4%; Score 73; DB 16; Length 14;	
Best Local Similarity	100.0%; Pred. No. 5.1e-06;	
Matches	14; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	2 VNYAFLHATDILLPA 15 	
DB	1 vnyafilhadtilla 14	
RESULT	4	

W18857
 ID W18857 standard; peptide: 14 AA.
 AC W18857;
 XX
 DT 05-JAN-1998 (first entry)
 XX
 DE 65 kD Glutamic acid decarboxylase peptide fragment 8.
 XX
 KW GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
 KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
 KW predilection: autoimmune; tumour; rheumatoid arthritis;
 KW multiple sclerosis.
 XX
 OS Synthetic.
 XX
 PN DE19526561-A1.
 XX
 PD 23-JAN-1997.
 XX
 PF 20-JUL-1995; 95DE-1026561.
 XX
 PR 20-JUL-1995; 95DE-1026561.
 XX
 PA (BOER) BOEHRINGER MANNHEIM GMBH.
 XX
 PI Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;
 PI Pozzilli P, Stahl P;
 XX
 DR WPI: 1997-088254/09.
 XX
 PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
 PT involving intradermal admin. of auto-reactive substances
 XX
 PS Claim 11: Fig 2; 12pp; German.
 XX
 SQ W18842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predilection to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 CC
 XX
 SQ Sequence 14 AA;
 XX

Query Match 66.4%; Score 73; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 5.1e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 VNYAFLLHATDLPLA 15
 |||||
 Db 1 vnyafllhacdllpa 14

RESULT 5
 R72267
 ID R72267 standard; Peptide: 20 AA.
 AC R72267;
 XX
 DT 13-NOV-1995 (first entry)
 XX
 DE Glutamic acid decarboxylase (GAD65) fragment.
 XX
 KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
 KW insulin-dependent diabetes mellitus; stiff man disease.
 XX
 OS Homo sapiens.
 XX

XX
 PN W09507992-A.
 XX
 PD 23-MAR-1995.
 XX
 PF 24-AUG-1994; 94WO-US09478.
 XX
 PR 17-SEP-1993; 93US-0123859.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
 PI WPI: 1995-131360/17.
 XX
 DR
 XX
 PT New polypeptide fragments of glutamic acid decarboxylase - for
 PT diagnosis and treatment of autoimmune disease, esp. insulin
 PT dependent diabetes, also related nucleic acid, vectors,
 PT antibodies, hybridoma(s) etc.
 XX
 PS Example 11, Page 76; 100pp; English.
 XX
 CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
 CC acid decarboxylase (GAD65) respectively, from which the GAD65
 CC fragments described in R72261-R72298 were derived. These fragments
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose
 CC and treat GAD-related autoimmune disorders, such as insulin
 CC dependant diabetes mellitus or stiff man disease.
 CC
 XX
 SQ Sequence 20 AA;
 XX

Query Match 66.4%; Score 73; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.6e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 8 HATDLPLACDGER 20
 |||||
 Db 1 hatdlplacdggr 13

RESULT 6
 Y59569
 ID Y59569 standard; peptide: 20 AA.
 AC Y59569;
 XX
 DT 03-APR-2000 (first entry)
 XX
 DE GAD65 fragment, peptide #7.
 XX
 KW GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
 KW insulin dependent diabetes mellitus; stiff man disease; diagnosis;
 KW therapy.
 XX
 OS Homo sapiens.
 XX
 PN US5998366-A.
 XX
 PD 07-DEC-1999.
 XX
 PF 09-APR-1997; 97US-0827618.
 XX
 PR 07-JUN-1995; 95US-0485725.
 PR 21-SEP-1990; 90US-0586536.
 PR 18-JUN-1991; 91US-0716909.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Tobin AJ, Kaufman DL, Erlander MG;
 PI WPI: 2000-095930/08.
 XX

PT Ameliorating glutamic acid decarboxylase associated autoimmune
 PT disorders such as insulin dependent diabetes mellitus and stiff man
 PT disease -
 XX
 PS Example 11; Column 42; 61pp; English.
 XX
 CC This sequence represents a fragment of the glutamic acid decarboxylase 65
 CC (GAD65) protein. The invention relates to a method of ameliorating GAD
 CC associated autoimmune disorders by administering a GAD65 peptide to the
 CC patient. The method can be used for ameliorating GAD associated
 CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
 CC and stiff man disease. GAD65 can also be useful for screening drugs that
 CC alter GAD function, for generating monoclonal antibodies and in
 CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
 CC and the diagnosis is quite easy. It is also possible to obtain much
 CC larger quantities of polypeptide via recombinant techniques than are
 CC available from natural sources.
 CC
 XX
 SQ Sequence 20 AA;

Query Match 66.4%; Score 73; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.6e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 HATDLPACDGER 20
 |||||||||
 Db 1 hatdlpacdger 13

RESULT 7

R72266
 ID R72266 standard; Peptide; 20 AA.

AC R72266;
 DT 13-NOV-1995 (first entry)
 XX

DE Glutamic acid decarboxylase (GAD65) fragment.

KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
 KW insulin-dependent diabetes mellitus; stiff man disease.

OS Homo sapiens.

PN W09507992-A.

PD 23-MAR-1995.

PF 24-AUG-1994; 94WO-US09478.

PR 17-SEP-1993; 93US-0123859.

XX (REGC) UNIV CALIFORNIA.

PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;

DR WPL; 1995-131360/17.

XX New polypeptide fragments of glutamic acid decarboxylase - for
 PT diagnosis and treatment of autoimmune disease, esp. insulin
 PT dependent diabetes, also related nucleic acid, vectors,
 PT antibodies, hybridoma(s) etc.
 XX

PS Claim 1; Page 76; 100pp; English.

XX O86481 and O86482 encode R71733 and R79105, rat and human glutamic
 CC acid decarboxylase (GAD65) respectively, from which the GAD65
 CC fragments described in R72261-R72298 were derived. These fragments
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose
 CC and treat GAD-related autoimmune disorders, such as insulin
 CC dependant diabetes mellitus or stiff man disease.
 XX

SQ Sequence 20 AA;

Query Match 58.2%; Score 64; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00025;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DVNYAFIHTDL 12
 |||||||||
 Db 9 dvnyafihadl 20

RESULT 8

Y59540
 ID Y59540 standard; peptide; 20 AA.

AC Y59540;
 XX
 DT 03-APR-2000 (first entry)
 XX

DE GAD65 fragment, peptide #6.

KW GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
 KW insulin dependent diabetes mellitus; stiff man disease; diagnosis;
 KW therapy.

OS Homo sapiens.

PN U55998366-A.

PD 07-DEC-1999.

PF 09-APR-1997; 97US-0827618.

PR 07-JUN-1995; 95US-0485725.

PR 21-SEP-1990; 90US-0586536.

PR 18-JUN-1991; 91US-0716909.

XX (REGC) UNIV CALIFORNIA.

PI Tobin AJ, Kaufman DL, Erlander MG;

DR WPL; 2000-095930/08.

XX Ameliorating glutamic acid decarboxylase associated autoimmune
 PT disorders such as insulin dependent diabetes mellitus and stiff man
 PT disease -
 XX

PS Claim 1; Column 33; 61pp; English.

XX This sequence represents a fragment of the glutamic acid decarboxylase 65
 CC (GAD65) protein. The invention relates to a method of ameliorating GAD
 CC associated autoimmune disorders by administering a GAD65 peptide to the
 CC patient. The method can be used for ameliorating GAD associated
 CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
 CC and stiff man disease. GAD65 can also be useful for screening drugs that
 CC alter GAD function, for generating monoclonal antibodies and in
 CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
 CC and the diagnosis is quite easy. It is also possible to obtain much
 CC larger quantities of polypeptide via recombinant techniques than are
 CC available from natural sources.
 XX

SQ Sequence 20 AA;

Query Match 58.2%; Score 64; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00025;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DVNYAFIHTDL 12
 |||||||||
 Db 9 dvnyafihadl 20

```

RESULT 9
ID Y59568 standard; peptide: 20 AA.
XX
AC Y59568;
XX
DT 03-APR-2000 (first entry)
XX
DE GAD65 fragment, peptide #6.
XX
KM GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
KM insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
KM therapy.
XX
OS Homo sapiens.
XX
PN US5998366-A.
XX
PD 07-DEC-1999.
XX
PE 09-APR-1997; 97US-0827618.
XX
PR 07-JUN-1995; 95US-0485725.
PR 21-SEP-1990; 90US-0586536.
PR 18-JUN-1991; 91US-0716909.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Tobin AJ, Kaufman DL, Erlander MG;
XX
DR WPI; 2000-095930/08.
XX
PT Ameliorating glutamic acid decarboxylase associated autoimmune
PT disorders such as insulin dependent diabetes mellitus and Stiff man
PT disease -
XX
PS Example 11; Column 42; 61pp; English.
XX
XX This sequence represents a fragment of the glutamic acid decarboxylase 65
CC (GAD65) protein. The invention relates to a method of ameliorating GAD
CC associated autoimmune disorders by administering a GAD65 peptide to the
CC patient. The method can be used for ameliorating GAD associated
CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
CC and Stiff man disease. GAD65 can also be useful for screening drugs that
CC alter GAD function, for generating monoclonal antibodies and in
CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
CC and the diagnosis is quite easy. It is also possible to obtain much
CC larger quantities of polypeptide via recombinant techniques than are
CC available from natural sources.
XX
SQ Sequence 20 AA:

```

Query Match 58.2%; Score 64; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00025;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 DNYAFELHATDL 12
   |||:|||||
DB 9 dnyafelhatdl 20

```

RESULT 10
 ID W79176 standard; peptide: 19 AA.
 XX
 AC W79176;
 XX
 DT 25-JAN-1999 (first entry)
 XX
 DE Fusion immunoglobulin heavy chain HIV gp120 T cell epitope #54.
 XX

```

KM B cell; T cell; epitope; immunoglobulin; heavy chain; gp120; IGH;
KM human immune deficiency virus; HIV; tolerance; treatment; therapy;
KM prophylaxis; vaccine; chemotherapy; immune response; modifier; tumour;
KM microbial infection; autoimmune disease; antibody; apoptosis;
KM antiviral T cell immunity.
XX
OS Mus sp.
XX
OS Homo sapiens.
XX
PN WO9836087-A1.
XX
PD 20-AUG-1998.
XX
PF 13-FEB-1998; 98WO-US02766.
XX
PR 13-FEB-1997; 97US-0040581.
XX
PA (AMNA-) AMERICAN NAT RED CROSS.
XX
PI Scott D, Zambidis E;
XX
DR WPI; 1998-506315/43.
XX
PT New fusion immunoglobulin heavy chain including gp120 epitopes and
PT related complete antibodies - DNA, vectors and transformed cells,
PT used to induce tolerance to the epitopes for treatment of human
PT immune deficiency virus infection
XX
PS Disclosure; Page 49; 154pp; English.
XX
XX This sequence is an epitope used in the construction of a novel fusion
CC immunoglobulin heavy chain (IGH) protein with a mammalian, especially
CC human, IGH chain fused in frame at its N-terminus to one or more human
CC immune deficiency virus (HIV) gp120 epitopes. Fusion Ig proteins and/or
CC transfected cells are used to tolerate subjects to gp120 epitopes and to
CC maintain this tolerance, particularly for treatment of HIV infection,
CC optionally together with other therapeutic/prophylactic agents such as
CC vaccines, chemotherapeutic agents and immune response modifiers. Such
CC proteins can be used against other diseases where an immune response is
CC deleterious, e.g. microbial infection, tumours or autoimmune disease.
CC Induction of tolerance suppresses production of antibodies against gp120,
CC so prevents or inhibits 'stander' apoptosis of uninfected T cells that
CC are bound to gp120 protein, maximising induction of protective antiviral
CC T cell immunity.
XX
SQ Sequence 19 AA:

```

Query Match 34.5%; Score 38; DB 19; Length 19;
 Best Local Similarity 42.9%; Pred. No. 5.4;
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

OY 4 YAFELHATDLAPAC 17
   |||:|:|:|
DB 6 yafelhatdlpald 19

```

RESULT 11
 ID R43901 standard; peptide: 15 AA.
 XX
 AC R43901;
 XX
 DT 06-JUN-1994 (first entry)
 XX
 DE HIV-1 IIIB gp120 monoclonal antibody antigenic peptide VI5P.
 XX
 KM Human immunodeficiency virus; therapy; HIV-1 infection; diagnosis;
 KM post-exposure prophylaxis; infection; cytotoxic agent targeting;
 KM neutralisation domain; BAT085; G3-136.
 XX
 OS Synthetic.
 XX

PN US5266478-A.
 XX
 PD 30-NOV-1993.
 XX
 PF 25-NOV-1991; 91US-0797692.
 XX
 PR 29-MAY-1987; 87US-0057445.
 PR 24-DEC-1987; 87US-0137861.
 PR 26-APR-1991; 91US-0692299.
 PR 26-SEP-1991; 91US-0767533.
 PR 25-NOV-1991; 91US-0797692.
 XX
 PA (TANO-) TANOX BIOSYSTEMS INC.
 XX
 PI Chang TW, Fung MSC, Sun BNC, Sun CRY;
 DR WPI; 1993-395339/49.
 XX
 PT Monoclonal antibody to HIV-1 gp120 - used for therapy of HIV-1
 PT Infection or AIDS, post-exposure prophylaxis and diagnosis of HIV-1
 PT HIV-1 infection
 XX
 PS Example; Page 8; 9pp; Japanese.
 XX
 CC The sequence is that of a synthetic peptide V15P corresponding to a
 CC unique neutralisation domain in the V2 region of HIV-1 gp 120. It
 CC can be used in the prodn. of monoclonal antibodies (Mabs) BAT085
 CC and G3-136. The Mabs may be used for therapy of HIV-1 infection or
 CC AIDS, post-exposure prophylaxis and diagnosis of HIV-1 infection and
 CC to reduce or eliminate virus infected T-cells by antibody-dependent
 CC cellular cytotoxicity, complement-mediated cytotoxicity or other
 CC cytolytic or regulatory immune mechanisms. The MAb can also be
 CC used for targeting cytotoxic agents to HIV-1 infected cells.
 CC
 CC Sequence 15 AA:
 SQ

Query Match 30.9%; Score 34; DB 14; Length 15;
 Best Local Similarity 45.5%; Pred. NO. 19;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 YAFPLHATDLP 14
 ||| : ||:
 Db 5 yaflykidlvp 15

RESULT 12
 W58542 W58542 standard; peptide; 15 AA.
 XX
 AC W58542;
 XX
 DT 02-SEP-1998 (first entry)
 XX
 DE Aspergillus oryzae phospholipase A1 peptide SEQ ID NO:3.
 XX
 KM Aspergillus oryzae; phospholipase A1; microbe; PLAI.
 XX
 OS Aspergillus oryzae.
 PN JP10155493-A.
 XX
 PD 16-JUN-1998.
 XX
 PF 03-OCT-1997; 97JP-0270967.
 XX
 PR 04-OCT-1996; 96JP-0264241.
 XX
 PA (SANY) SANKYO CO LTD.
 XX
 DR WPI; 1998-391046/34.
 XX
 PT Aspergillus-derived phospholipase A1 gene - used for the recombinant

PT production of phospholipase A of high purity and in a high yield
 XX
 XX Example 1; Page 18; 23pp; Japanese.
 PS
 XX
 CC The present sequence represents a phospholipase A1 (PLAI) peptide
 CC derived from Aspergillus oryzae. Phospholipase A can be prepared by
 CC culturing host cells and collecting the host cell from the culture. DNA
 CC encoding mature PLAI can be used to produce recombinant phospholipase
 CC A of high purity and in a high yield.
 XX
 XX Sequence 15 AA:
 SQ

Query Match 30.9%; Score 34; DB 19; Length 15;
 Best Local Similarity 41.7%; Pred. NO. 19;
 Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 NYAFPLHATDLP 14
 ||| : ||:
 Db 3 nyrvthndilvp 14

RESULT 13
 W87621 W87621 standard; peptide; 15 AA.
 XX
 AC W87621;
 XX
 DT 03-MAR-1999 (first entry)
 XX
 DE Epitope of HIV-1 gp120 protein which binds antibody BAT085.
 XX
 KM Epitope; gp120 protein; monoclonal antibody; HIV-1; antibody BAT123;
 KM antibody BAT267; antibody BAT085; T cell infection inhibition;
 KM syncytia formation; acquired immune deficiency syndrome; AIDS;
 KM AIDS-related complex; passive immunisation; antiviral; cytotoxic;
 KM viral load measurement; vaccine.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN US5854400-A.
 XX
 PD 29-DEC-1998.
 XX
 PF 22-SEP-1992; 92US-0950571.
 XX
 PR 22-SEP-1992; 92US-0950571.
 PR 29-MAY-1987; 87US-0057445.
 PR 27-DEC-1987; 87US-0137861.
 PR 26-SEP-1991; 91US-0767533.
 XX
 PA (TANO-) TANOX INC.
 XX
 PI Chang NT, Chang TW, Fung MSC, Sun BNC, Sun CRY;
 DR WPI; 1999-095002/08.
 XX
 PT Monoclonal antibodies directed against regions of gp120 of human
 PT immune deficiency virus-1 - are neutralising and able to inhibit
 PT infection of T cells and formation of syncytia, used for treatment,
 PT prevention or diagnosis of acquired immune deficiency syndrome
 XX
 PS Claim 8; Column 9; 16pp; English.
 XX
 CC The present sequence represents an epitope of the gp120 protein of
 CC human immune deficiency virus (HIV)-1. The sequence comprises
 CC amino acids 169 to 183 of gp120. The specification describes
 CC monoclonal antibodies which bind to epitopes of the gp120 protein.
 CC Specifically, these antibodies are designated BAT123, 267 and 085.
 CC Monoclonal antibodies neutralise HIV-1, inhibiting both infection
 CC of T cells and formation of syncytia, so are used to treat acquired
 CC immune deficiency syndrome (AIDS) and AIDS-related complex, by
 CC passive immunisation, as carriers of cytotoxic or antiviral agents,

CC and in extracorporeal systems. They can also be used as immunoassay
 CC reagents (for diagnosis or measurement of viral load) and to screen
 CC for neutralising epitopes, potentially useful in vaccine development.
 XX

SO Sequence 15 AA;

Query Match 30.9%; Score 34; DB 20; Length 15;
 Best Local Similarity 45.5%; Pred. No. 19;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 YAFPHATDLP 14
 |||:|:|:
 Db 5 yafyfxdlip 15

RESULT 14

W72822
 ID W72822 standard; peptide; 15 AA.

AC W72822;

DT 13-JAN-1999 (first entry)

DE HIV-1 gp120 monoclonal antibody BAT085 residue 178 to 192.

KW HIV-1; gp120; epitope; monoclonal antibody; envelope; neutralise;
 KW inhibit; infection; T-cell; inhibit syncytium formation; AIDS.

OS Human immunodeficiency virus type 1.

PN US5834599-A.

PD 10-NOV-1998.

PF 04-MAR-1993; 93US-0026276.

PR 04-MAR-1993; 93US-0026276.

PR 29-MAR-1987; 87US-0057445.

PR 24-DEC-1987; 87US-0137861.

PR 25-APR-1989; 89US-0343540.

PR 05-JUN-1992; 92US-0895197.

PA (TANO-) TANOX BIOSYSTEMS INC.

PI Chang NT, Chang T, Fung SC, Kim YW, Sun BN, Sun CR;

DR WPI; 1999-008810/01.

XX Antibody conjugate comprising monoclonal antibody - which binds to
 PT epitope within amino acid residue of gp120 which neutralises HIV-1
 PT conjugated with, e.g. cytotoxic agent

PS Example 4; Column 25; 22pp; English.

XX The present invention describes an antibody conjugate comprising an
 CC antibody (Ab) which binds to an epitope within amino acid residue
 CC 308-322 of gp120 and neutralises HIV-1, conjugated with a cytotoxic
 CC agent, an anti-viral agent or an agent which facilitates passage through
 CC the blood brain barrier. Also described is an antibody conjugate as
 CC above but where the Ab binds to an epitope within amino acid residue
 CC 288-312 of gp120 which neutralises HIV-1. The present sequence represents
 CC an HIV-1 gp120 monoclonal antibody BAT085 residue 178 to 192 from an
 CC example of the present invention. The Ab are monoclonal Ab which bind
 CC to the gp120 protein on the envelope of HIV-1. They inhibit the
 CC infection of T-cells and also inhibit syncytium formation. The
 CC antibodies are group specific and neutralise different strains and
 CC isolates of HIV-1. The antibodies have a variety of uses, including the
 CC treatment and prevention of AIDS and AIDS related complex. They are
 CC especially used to kill infected T-cells.

SO Sequence 15 AA;

Query Match 30.9%; Score 34; DB 20; Length 15;
 Best Local Similarity 45.5%; Pred. No. 19;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 YAFPHATDLP 14
 |||:|:|:
 Db 5 yafyfxdlip 15

RESULT 15

W76928
 ID W76928 standard; peptide; 20 AA.

AC W76928;

DT 25-JAN-1999 (first entry)

DE Fusion immunoglobulin heavy chain HIV gp120 B cell epitope #68.

KW B cell; T cell; epitope; immunoglobulin; heavy chain; gp120; IgH;

KW human immune deficiency virus; HIV; tolerance; treatment; therapy;

KW prophylaxis; vaccine; chemotherapy; immune response; modifier; tumour;

KW microbial infection; autoimmune disease; antibody; apoptosis;

OS anti-viral T cell immunity.

OS Mus sp.

OS Homo sapiens.

PN WO9836087-A1.

PD 20-AUG-1998.

PF 13-FEB-1998; 98WO-US02766.

PR 13-FEB-1997; 97US-0040581.

PA (AMNA-) AMERICAN NAT RED CROSS.

PI Scott D, Zambidis E;

DR WPI; 1998-506315/43.

XX New fusion immunoglobulin heavy chain including gp120 epitopes and
 PT related complete antibodies - DNA, vectors and transformed cells,
 PT used to induce tolerance to the epitopes for treatment of human
 PT immune deficiency virus infection

PS Disclosure; Page 38; 154pp; English.

XX This sequence is an epitope used in the construction of a novel fusion
 CC immunoglobulin heavy chain (IGH) protein with a mammalian, especially
 CC human, IGH chain fused in frame at its N-terminus to one or more human
 CC immune deficiency virus (HIV) gp120 epitopes. Fusion Ig proteins and/or
 CC transfected cells are used to tolerate subjects to gp120 epitopes and to
 CC maintain this tolerance, particularly for treatment of HIV infection,
 CC optionally together with other therapeutic/prophylactic agents such as
 CC vaccines, chemotherapeutic agents and immune response modifiers. Such
 CC proteins can be used against other diseases where an immune response is
 CC deleterious, e.g. microbial infection, tumours or autoimmune disease.
 CC Induction of tolerance suppresses production of antibodies against gp120,
 CC so prevents or inhibits 'bystander' apoptosis of uninfected T cells that
 CC are bound to gp120 protein, maximising induction of protective antiviral
 CC T cell immunity.

SO Sequence 20 AA;

SO Sequence 20 AA;

Query Match 30.9%; Score 34; DB 19; Length 20;
 Best Local Similarity 45.5%; Pred. No. 27;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 YAFPHATDLP 14

Db 4 ihtae11aac 13

RESULT 18
P82057
ID P82057 standard: protein; 20 AA.
XX
AC P82057;
XX
XX
DT 22-OCT-1990 (first entry)
XX
DE Pep-14 comprising amino acids 195-214 of Factor VIII binding peptide.
XX
KW Glycoprotein Ib (GP1b) Factor VIII binding region;
XX von Willebrand Factor (VWF); factor VIII:C-VWF complexes.
XX
OS synthetic.
XX
PN EP295645-A.
XX
PD 21-DEC-1988.
XX
PF 15-JUN-1988; 88EP-0109532.
XX
PR 16-JUN-1987; 87US-0062896.
PR 02-MAR-1988; 88US-0162877.
XX
PA (ZYMO-) ZYMOGENETICS INC.
XX
PI Kumar AA, Hagen FS, Sledziewski AZ;
XX
DR WPI; 1988-362157/51.
XX
PT New blood factor peptides -
PT Specifically binding to von Willebrand factor or factor VIII:C
XX
PS Example; Page 7; 12pp; English.
XX
CC The region of GP1b found to be responsible for its ability to bind
CC to vWF lies between amino acids 165-260. Overlapping peptides derived
CC from this region and having terminal Cys or Lys residues are found to
CC bind specifically to vWF. These peptides can be used to isolate vWF or
CC factor VIII:C-vWF complex.
CC See also P82055-6 or P82058-P82060.
XX
SQ Sequence 20 AA;

Query Match 29.1%; Score 32; DB 9; Length 20;
Best Local Similarity 42.9%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 4 YAFHARDLPACD 17
:||||| 1 1:
5 fafhgprw1nce 18

RESULT 19
R23839
ID R23839 standard: peptide; 17 AA.
XX
AC R23839;
XX
DT 05-NOV-1992 (first entry)
XX
DE Peptide from the ED1 domain of fibronectin.
XX
KW FN; toxemia; pregnancy; type III.
XX
OS Homo sapiens.
XX
PN US5108898-A.
XX

PD 28-APR-1992.
XX
XX 18-JAN-1989; 89US-0298622.
XX
PR 18-JAN-1989; 89US-0298622.
XX
PA (PETE/) PETERS J H.
XX
XX
PI Lockwood CJ, Peters JH;
XX
DR WPI; 1992-166519/20.
XX
XX
PT Predicting toxemia in pregnancy - by detecting elevated levels
PT of fibronectin having variably included type III repeat region in
PT body fluid sample
XX
PS Claim 10; Page 8; 9pp; English.

CC The peptide was synthesised by standard methods and has a sequence
CC based on residues 45-60 of the ED1 sequence. The peptide was used
CC to raise antibodies which will immunoreact with ED1 but not with a
CC plasma fibronectin monomer. The antibodies can be used for detecting
CC fibronectin contg. variably included type III repeats in a sample,
CC thus determining patients destined to develop toxemia, partic.
CC preclampsia, prior to the onset of maternal signs and symptoms
CC of the disease, allowing for early therapeutic intervention.
CC See also R23837-42.
XX
SQ Sequence 17 AA;

Query Match 28.2%; Score 31; DB 13; Length 17;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 11 DLPACDGE 19
:||||| 1 1:
1 elfpdpge 9

RESULT 20
R23840
ID R23840 standard: peptide; 19 AA.
XX
AC R23840;
XX
DT 05-NOV-1992 (first entry)
XX
DE Peptide from the ED1 domain of fibronectin.
XX
KW FN; toxemia; pregnancy; type III.
XX
OS Homo sapiens.
XX
PN US5108898-A.
XX
PD 28-APR-1992.
XX
PF 18-JAN-1989; 89US-0298622.
XX
PR 18-JAN-1989; 89US-0298622.
XX
PA (PETE/) PETERS J H.
XX
PI Lockwood CJ, Peters JH;
XX
DR WPI; 1992-166519/20.
XX
XX
PT Predicting toxemia in pregnancy - by detecting elevated levels
PT of fibronectin having variably included type III repeat region in
PT body fluid sample
XX
PS Claim 10; Page 8; 9pp; English.

XX The peptide was synthesised by standard methods and has a sequence
 CC based on residues 45-60 of the ED1 sequence. The peptide was used
 CC to raise antibodies which will immunoreact with ED1 but not with a
 CC plasma fibronectin monomer. The antibodies can be used for detecting
 CC fibronectin conrg. variably included Type III repeats in a sample,
 CC thus determining patients destined to develop toxemia, partic.
 CC preeclampsia, prior to the onset of maternal signs and symptoms
 CC of the disease, allowing for early therapeutic intervention.
 CC See also R23837-42.

XX Sequence 19 AA;

Query Match 28.2%; Score 31; DB 13; Length 19;
 Best Local Similarity 66.7%; Pred. No. 80;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 11 DLPACDGE 19
 : | | | | |
 Db 1 elifpdg 9

RESULT 21

W96277
 ID W96277 standard; Peptide; 21 AA.

AC W96277;

DT 14-JUN-1999 (first entry)

DE B-cell epitope of testis specific isoform of calpastin.

XX Calpastin; isoform; testis; antibodies; vaccine; inhibition;
 KW fertilisation; egg; spermatazoa; sperm; contraceptive;

KW B-cell epitope; T-cell epitope.

XX Homo sapiens.

PN W09903490-A1.

PD 28-JAN-1999.

PE 21-JUL-1998; 98WO-US15094.

PR 20-JUL-1998; 98US-0119149.

XX 21-JUL-1997; 97US-0053337.

PA (NOUN) UNIV NORTHWESTERN.

PI Goldberg E;

XX WPI; 1999-131861/11.

PT New testis-specific isoforms of calpastatin peptides - used in
 PT vaccines and for the production of antibodies which can inhibit the
 PT fertilisation of an egg by sperm in a mammal

PS Claim 8; Page 34; 48pp; English.

XX Testis-specific calpastatin and peptides derived from it can be used
 CC for detecting antibodies for assessing infertility in a patient. The
 CC peptides and antibodies raised by them can also be used for
 CC inhibiting fertilisation of an egg by a sperm in a mammal.
 CC Testis specific calpastatin peptides can therefore be used in the
 CC production of a vaccine used to inhibit fertilisation of an egg by
 CC sperm. The vaccine can be used as topical contraceptives in mammals,
 CC including dogs, cats and other domestic animals and humans.

XX Sequence 21 AA;

Query Match 28.2%; Score 31; DB 20; Length 21;

Best Local Similarity 100.0%; Pred. No. 90;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 CDGER 20
 : | | | | |
 Db 3 cdger 7

RESULT 22

W30416
 ID W30416 standard; peptide; 7 AA.

AC W30416;

DT 22-APR-1998 (first entry)

DE HRE-I affinity peptide SEQ ID NO:34.

XX Binding site; HRE-I; screening; H-Ras; target gene; plasmid;
 KW inhibition; affinity peptide.

XX Synthetic.

PN W09737030-A1.

PD 09-OCT-1997.

PE 31-MAR-1997; 97WO-JP01105.

PR 03-OCT-1996; 96JP-0263345.

XX 01-APR-1996; 96JP-0101990.

PA (YAWH) NIPPON STEEL CHEM CO.

XX (YAWA) NIPPON STEEL CORP.

PI Kato T;

DR WPI; 1997-503115/46.

PT Plasmid for screening peptide(s) which bind target genes - for
 PT treatment of conditions associated with H-Ras

PS Claim 41; Page 69; 127pp; Japanese.

XX A novel plasmid has been developed for screening peptides which bind
 CC target genes. The plasmid comprises a sequence encoding the test
 CC peptide, a promoter comprising the target gene, and a reporter gene
 CC bound to the promoter. The present sequence represents a specifically
 CC claimed peptide with affinity to HRE-I. The plasmids are used for
 CC screening for peptides which bind to target genes. The identified
 CC peptides can be used for the treatment of conditions associated with
 CC the inhibition of the expression the genes, and treatment of conditions
 CC associated with H-Ras.

XX Sequence 7 AA;

Query Match 27.3%; Score 30; DB 16; Length 7;
 Best Local Similarity 71.4%; Pred. No. 2; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 2;

OY 14 PACDGER 20
 : | | | | |
 Db 1 pycdker 7

RESULT 23

R89429
 ID R89429 standard; Protein; 15 AA.

AC R89429;

DT 13-SEP-1996 (first entry)

XX JP06199894-A.
PN
XX 19-JUL-1994.
PD
XX
XX 02-APR-1993; 93JP-0076791.
PE
XX 27-AUG-1992; 92JP-0228965.
PR 10-NOV-1992; 92JP-0299691.
XX
XX (ASAH) ASAH KASEI KOGYO KK.
PA (INOM/) INOMAMARI M.
XX
XX WPI; 1994-269451/33.
DR
XX T cell epitope present in the core protein region of Hepatitis C
PT virus (HCV) - used for activation of cellular immunity mechanisms
PT
XX
XX Claim 5; Fig 2; 14pp; Japanese.
PS
XX A T-cell stimulating peptide is claimed which is ca. 5-20 amino
CC acids long and part of the core protein of HCV and is recognised by
CC and stimulates T-cells. A CD8-positive T-cell stimulating peptide
CC has the sequence given in R56606 (NP-9) or R56616 (9MA). A CD4-
CC positive T-cell stimulating peptide has the sequence given in R56609
CC (NP-12) or R56614 (NP-17). Synthetic peptide mixts. (Mix A: NP-1 -
CC NP-5, Mix B: NP-6 - NP-10, Mix C: NP-11 - NP14, MixD: NP-15 - NP-18)
CC were used in experiments to evaluate activity.
CC Peptides 9MA, 9CM and 9CA (R56616-18) are variants of peptide NP-9.
XX
SQ Sequence 20 AA;

Query Match 27.3%; Score 30; DB 15; Length 20;
Best Local Similarity 53.3%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 4; Gaps 1;
OY 2 VNVAFLHATDLPAC 16
||| || |||
Db 2 vny-----atgnlpgc 12

Search completed: February 5, 2001, 10:47:29
Job time: 623 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:49:26 ; Search time 58.67 Seconds
(without alignments)
23.147 Million cell updates/sec

Title: US-08-981-824-2

Perfect score: 102

Sequence: 1 SNMTAMMIARFKMPPEVK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 5001

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR66:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	24.5	14	2	PI0142
2	25	24.5	25	2	S11383
3	24	23.5	25	2	A34921
4	23	22.5	10	2	A60476
5	23	22.5	13	2	PN0122
6	23	22.5	14	2	PH1614
7	23	22.5	17	2	PH1617
8	23	22.5	17	2	I34269
9	23	22.5	22	2	S63674
10	23	22.5	22	2	S03263
11	22	21.6	15	2	A48372
12	22	21.6	19	2	S03076
13	22	21.6	20	2	A54083
14	22	21.6	20	2	C36894
15	22	21.6	20	2	A56894
16	22	21.6	20	2	PC7073
17	21	20.6	9	2	T09741
18	21	20.6	16	2	S32587
19	21	20.6	17	2	S14661
20	21	20.6	18	2	S02808
21	21	20.6	19	2	S57286
22	21	20.6	20	2	S51066
23	21	20.6	21	2	A39108
24	21	20.6	25	2	S68642
25	21	20.6	25	2	S68642
26	21	20.6	25	2	S68642
27	20.5	20.1	25	2	S17683
28	20	19.6	8	2	A61597
29	20	19.6	13	2	S22689

30	20	19.6	14	2	S36578	dodecenyl-CoA Del
31	20	19.6	16	2	S65709	major allergen Myr
32	20	19.6	17	2	S78421	ribosomal protein
33	20	19.6	18	2	S29264	ovohemerythrin - d
34	20	19.6	21	2	PQ0145	glucan endo-1,3-be
35	20	19.6	21	2	S78416	ribosomal protein
36	20	19.6	22	2	S13977	chlorophyll a/b-bl
37	20	19.6	23	2	D60583	glycoprotein horio
38	20	19.6	23	4	JE0016	probable 2.9k prot
39	20	19.6	24	2	S38729	probable malate ca
40	20	19.6	24	2	S53793	actin - mouse (fra
41	20	19.6	25	2	S09540	gene 1.5 protein -
42	19.5	19.1	16	2	F44908	chitinase (EC 3.2.
43	19	18.6	11	2	G42762	multicatalytic end
44	19	18.6	12	2	PH1587	Ig H chain V-D-J r
45	19	18.6	14	2	A61308	hemocyanin chain 2

ALIGNMENTS

RESULT 1
PI0142
carbon-monoxide dehydrogenase (EC 1.2.99.2) medium chain - pseudomonas carboxydoflava
C:Species: Pseudomonas carboxydoflava
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Apr-1993
C:Accession: PI0142
R:Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.
Arch. Microbiol. 152, 335-341, 1989
A>Title: Homology and distribution of CO dehydrogenase structural genes in carboxydot
A:Reference number: PI0138; MUID:90055678
A:Accession: PI0142
A:Molecule type: protein
A:Residues: 1-14 <KRA>
A:Comment: Carbon-monoxide dehydrogenase consists of three polypeptide chains: large,
C:Keywords: oxidoreductase

Query Match 24.5% Score 25; DB 2; Length 14;
Best Local Similarity 71.4% Pred. No. 4.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 MMARPK 12
DB 1 MMIPRF 7

RESULT 2
S11383
uridine phosphorylase (EC 2.4.2.3) - Lactobacillus casei (fragment)
N:Alternate names: pyrimidine phosphorylase
C:Species: Lactobacillus casei
C>Date: 19-Mar-1997 #sequence_revision 26-Feb-1998 #text_change 26-Feb-1998
C:Accession: S11383
R:Avraham, Y.; Grossowicz, N.; Yashphe, J.
Biochim. Biophys. Acta 1040, 287-293, 1990
A>Title: Purification and characterization of uridine and thymidine phosphorylase fro
A:Reference number: S11383; MUID:90381286
A:Accession: S11383
A:Molecule type: protein
A:Residues: 1-25 <AVR>
A:Experimental source: strain ATCC 7469
C:Keywords: glycosyltransferase; homotetramer; pentosyltransferase

Query Match 24.5% Score 25; DB 2; Length 25;
Best Local Similarity 55.6% Pred. No. 8.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 12 KMPEVK 20
DB 4 KVIPETK 12

RESULT 3
A34921
ribulose-bisphosphate carboxylase (EC 4.1.1.39) small chain - Flaveria linearis (fragment)
C:Species: Flaveria linearis
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 12-Apr-1995
C:Accession: A34921
R:Hudson, G.S.; Mahon, J.D.; Anderson, P.A.; Gibbs, M.J.; Badger, M.R.; Andrews, T.J.; W
J. Biol. Chem. 265, 808-814, 1990
A:Title: Comparisons of rbcL genes for the large subunit of ribulose-bisphosphate carbox
A:Reference number: A34921; MUID:90110139
A:Accession: A34921
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <HUD>
C:Superfamily: ribulose-bisphosphate carboxylase small chain
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 23.5%; Score 24; DB 2; Length 25;
Best Local Similarity 40.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 11 FKMPPEVKEK 20
|::|::|
Db 1 FKWMPPLCKK 10

RESULT 4
A60476
S-layer protein - Bacillus thuringiensis (fragment)
C:Species: Bacillus thuringiensis
C:Date: 20-Feb-1993 #sequence_revision 20-Feb-1993 #text_change 12-Mar-1993
C:Accession: A60476
R:Luckevich, M.D.; Beveridge, T.J.
J. Bacteriol. 171, 6656-6667, 1989
A:Title: Characterization of a dynamic S layer on Bacillus thuringiensis.
A:Reference number: A60476; MUID:90078111
A:Accession: A60476
A:Molecule type: protein
A:Residues: 1-10 <LUC>
C:Comment: The S-layer, or surface array, is the outermost component of several archaea

Query Match 22.5%; Score 23; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 7e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 12 KMPEV 17
|::|::|
Db 3 KTFPDV 8

RESULT 5
PN0122
OIL protein - vaccinia virus (strain L-IVP) (fragment)
C:Species: vaccinia virus
C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 20-Jun-2000
C:Accession: PN0122
R:Razankina, O.I.; Shchelkunov, S.N.; Muravlev, A.I.; Netesova, N.A.; Mikhajukov, N.N.;
Mol. Biol. (Mosk.) 24, 968-976, 1990
A:Title: The molecular biological study of vaccinia virus genome II: localization and f
A:Reference number: PN0119; MUID:91066899
A:Accession: PN0122
A:Molecule type: DNA
A:Residues: 1-13 <RU>
C:Superfamily: variola major virus hypothetical protein Q1L

Query Match 22.5%; Score 23; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 9.1e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 13 MPEVKEK 20
|::|::|
Db 3 MYPEFANK 10

RESULT 6
PH1614
Ig H chain V-D-J region (clone B-less 18) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1614
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less m
A:Reference number: PH1580; MUID:93301609
A:Accession: PH1614
A:Molecule type: DNA
A:Residues: 1-14 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match 22.5%; Score 23; DB 2; Length 14;
Best Local Similarity 80.0%; Pred. No. 9.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 NMYAM 6
|::|::|
Db 8 NMYAM 12

RESULT 7
PH1617
Ig H chain V-D-J region (clone B-less 32) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1617
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less m
A:Reference number: PH1580; MUID:93301609
A:Accession: PH1617
A:Molecule type: DNA
A:Residues: 1-14 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match 22.5%; Score 23; DB 2; Length 14;
Best Local Similarity 80.0%; Pred. No. 9.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 NMYAM 6
|::|::|
Db 8 NMYAM 12

RESULT 8
I54269
vitamin D binding protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000
C:Accession: I54269
R:Brann, A.; Bichlmayer, R.; Muller, B.; Cleve, H.
Hum. Genet. 90, 526-532, 1993
A:Title: Molecular evaluation of an Alu repeat including a polymorphic variable poly
A:Reference number: I54269; MUID:93154720
A:Accession: I54269
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-17 <RES>
A:Cross-references: GB:S54074; NID:g264876; PIDN:AAD13872.1; PID:g4261572
C:Superfamily: serum albumin; serum albumin repeat homology

Query Match 22.5%; Score 23; DB 2; Length 17;
 Best Local Similarity 35.7%; Pred. No. 1.2e+03;
 Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 4 YAMMIRFKMPEV 17
 | : : |
 Db 1 YTFELSRRTLPEV 14

RESULT 9

S63674
 28 K secretory protein - rat (fragments)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
 C:Accession: S63674
 R:Pesheko, I.V.; Novoselov, V.I.; Evdokimov, V.A.; Nikolaev, Y.V.; Shuvaeva, T.M.; Lipy
 FEBS Lett. 381, 12-14, 1996
 A:Title: Novel 28-kDa secretory protein from rat olfactory epithelium.
 A:Reference number: S63674; MUID:96193902
 A:Accession: S63674
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-5,6-11,12-17,18-22 <PES>

Query Match 22.5%; Score 23; DB 2; Length 22;
 Best Local Similarity 50.0%; Pred. No. 1.6e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 12 KMPEVKE 19
 | : : |
 Db 6 KLAPEKFD 13

RESULT 10

S06263
 gastrin-releasing peptide - smaller spotted catshark (tentative sequence) (fragment)
 C:Species: Scyliorhinus canicula (smaller spotted catshark, smaller spotted dogfish)
 C:Date: 31-Mar-1990 #sequence_revision 30-Jan-1998 #text_change 31-Mar-2000
 C:Accession: S06263
 R:Conlon, J.M.; Henderson, I.W.; Thim, L.
 Gen. Comp. Endocrinol. 68, 415-420, 1987
 A:Title: Gastrin-releasing peptide from the intestine of the elasmobranch fish, scyliorhin
 A:Reference number: S06263; MUID:88137922
 A:Accession: S06263
 A:Molecule type: protein
 A:Residues: 1-25 <CON>
 A:Note: the sequence from the summary is inconsistent with that from table 1 and table 3
 C:Superfamily: gastrin-releasing peptide
 C:Keywords: neuropeptide

Query Match 22.5%; Score 23; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 KMFP 15
 | : : |
 Db 11 KMFP 14

RESULT 11

A48372
 benzoyl-CoA ligase - Methanospirillum hungatei (fragment)
 C:Species: Methanospirillum hungatei
 C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 07-May-1999
 C:Accession: A48372
 R:Auburger, G.; Winter, J.
 Appl. Microbiol. Biotechnol. 37, 789-795, 1992
 A:Title: Purification and characterization of benzoyl-CoA ligase from a syntrophic, benz
 A:Reference number: A48372; MUID:93040109

A:Accession: A48372
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-15 <AUB>
 A:Note: sequence extracted from NCBI backbone (NCBIP:118357)

Query Match 21.6%; Score 22; DB 2; Length 15;
 Best Local Similarity 60.0%; Pred. No. 1.6e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 KMPE 16
 | : : |
 Db 2 KLYPE 6

RESULT 12

S03076
 T-cell receptor gamma chain J region (JPI) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 30-May-1997
 C:Accession: S03076
 R:Huck, S.; Lefranc, M.P.
 FEBS Lett. 224, 291-296, 1987
 A:Title: Rearrangements to the JPI, JP and JP2 segments in the human T-cell rearrangl
 A:Reference number: S03076; MUID:88083552
 A:Accession: S03076
 A:Molecule type: DNA
 A:Residues: 1-19 <HUC>
 A:Cross-references: EMBL:X08084
 C:Keywords: T-cell receptor

Query Match 21.6%; Score 22; DB 2; Length 19;
 Best Local Similarity 66.7%; Pred. No. 2e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 FMPEPE 16
 | : : |
 Db 5 FKFAE 10

RESULT 13

AS4083
 p190/210, fatty acid synthase, p140ex2 strand exchange protein activator - fission y
 C:Species: Schizosaccharomyces pombe
 C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
 C:Accession: AS4083
 R:Kaaslan, E.; Heyer, W.D.
 J. Biol. Chem. 269, 14103-14110, 1994
 A:Title: Schizosaccharomyces pombe fatty acid synthase mediates DNA strand exchange 1
 A:Reference number: AS4083; MUID:94245730
 A:Accession: AS4083
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-20 <KAS>
 A:Note: sequence extracted from NCBI backbone (NCBIP:148744)
 C:Superfamily: yeast fatty-acid synthase

Query Match 21.6%; Score 22; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 2.1e+03;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 13 MPEVKEK 20
 | : : |
 Db 1 MPEVKE 8

RESULT 14

C56894
 Intracytoplasmic chromoprotein 1 - Neothyris lenticularis (fragment)
 C:Species: Neothyris lenticularis

C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 20-Jun-2000
 C/Accession: C56894
 R/Cusack, M.; Curry, G.; Clegg, H.; Abbott, G.
 Comp. Biochem. Physiol. B 102, 93-95, 1992
 A>Title: An intracrystalline chromoprotein from red brachiopod shells: implications for
 A/Reference number: A56894; MUID:92405551
 A/Contents: Deshayes, red brachiopod shells
 A/Accession: C56894
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 1-20 <CUS>
 A/Note: sequence extracted from NCBI backbone (NCBIP:114884)
 C/Keywords: Chromoprotein

Query Match 21.6%; Score 22; DB 2; Length 20;
 Best Local Similarity 57.1%; Pred. No. 2.1e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 YAMMIAR 10
 |||:
 Db 7 YATMISK 13

RESULT 15
 A56894
 Intracrystalline chromoprotein I - Terebratella sanguinea (fragment)
 C/Species: Terebratella sanguinea
 C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 20-Jun-2000
 C/Accession: A56894
 R/Cusack, M.; Curry, G.; Clegg, H.; Abbott, G.
 Comp. Biochem. Physiol. B 102, 93-95, 1992
 A>Title: An intracrystalline chromoprotein from red brachiopod shells: implications for
 A/Reference number: A56894; MUID:92405551
 A/Contents: Leach, red brachiopod shells
 A/Accession: A56894
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 1-20 <CUS>
 A/Note: sequence extracted from NCBI backbone (NCBIP:114882)
 C/Keywords: Chromoprotein

Query Match 21.6%; Score 22; DB 2; Length 20;
 Best Local Similarity 57.1%; Pred. No. 2.1e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 YAMMIAR 10
 |||:
 Db 7 YATMISK 13

RESULT 16
 B56894
 Intracrystalline chromoprotein I - Waltonia inconspicua (fragment)
 C/Species: Waltonia inconspicua
 C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 20-Jun-2000
 C/Accession: B56894
 R/Cusack, M.; Curry, G.; Clegg, H.; Abbott, G.
 Comp. Biochem. Physiol. B 102, 93-95, 1992
 A>Title: An intracrystalline chromoprotein from red brachiopod shells: implications for
 A/Reference number: A56894; MUID:92405551
 A/Contents: Sowerby, red brachiopod shells
 A/Accession: B56894
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 1-20 <CUS>
 A/Note: sequence extracted from NCBI backbone (NCBIP:114883)
 C/Keywords: Chromoprotein

Query Match 21.6%; Score 22; DB 2; Length 20;
 Best Local Similarity 57.1%; Pred. No. 2.1e+03;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 4 YAMMIAR 10
 |||:
 Db 7 YATMISK 13

RESULT 17
 PC7073
 ubiquinol--cytochrome-c reductase (EC 1.10.2.2) core protein II - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
 C/Accession: PC7073
 R/Isugita, A.; Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Matsui, T.; Watanabe, Y
 Electrophoresis 21, 1853-1871, 2000
 A>Title: Proteome analysis of mouse brain: Two-dimensional electrophoresis profiles o
 A/Reference number: PC7072
 A/Accession: PC7073
 A/Molecule type: protein
 A/Residues: 1-9 <TSU>
 C/Keywords: brain; core protein; oxidoreductase

Query Match 20.6%; Score 21; DB 2; Length 9;
 Best Local Similarity 57.1%; Pred. No. 2e+05;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 12 KMPEVK 18
 |||:
 Db 3 KVAPEVK 9

RESULT 18
 T09741
 photosystem I chain psal - upland cotton chloroplast (fragment)
 C/Species: chloroplast Gossypium hirsutum (upland cotton)
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C/Accession: T09741
 R/Small, R.L.; Ryburn, J.A.; Cronm, R.C.; Seelanan, T.; Wendel, J.F.
 Am. J. Bot. 85, 1301-1315, 1998
 A>Title: The tortoise and the hare: choosing between noncoding plastome and nuclear A
 A/Reference number: Z16323
 A/Accession: T09741
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-16 <SMA>
 A/Cross-references: EMBL:AF031581; NID:92623684; PID:93723945
 C/Genetics:
 A/Genes: psal
 A/Genome: psal
 C/Keywords: chloroplast; photosynthesis; photosystem I

Query Match 20.6%; Score 21; DB 2; Length 16;
 Best Local Similarity 42.9%; Pred. No. 2.5e+03;
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 11 FKPEPV 17
 |||:
 Db 4 FSSFPST 10

RESULT 19
 S32587
 L-ascorbate peroxidase (EC 1.11.1.11) isozyme II - spinach (fragment)
 C/Species: Spinacia oleracea (spinach)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-Sep-1993
 C/Accession: S32587; S15878
 R/Kudo, A.; Saij, H.; Tanaka, K.; Tanaka, K.; Kondo, N.
 Plant Mol. Biol. 18, 691-701, 1992
 A>Title: Cloning and sequencing of a cDNA encoding ascorbate peroxidase from Arabidop
 A/Reference number: S20866; MUID:92216045
 A/Accession: S32587

A:Molecule type: protein
 A:Residues: 1-17 <RUB>
 A>Note: This is a revision to the sequence from reference S15878
 R.Tanaka, K.; Takeuchi, E.; Kubo, A.; Sakaki, T.; Haraguchi, K.; Kawamura, Y.
 Arch. Biochem. Biophys. 286, 371-375, 1991
 A>Title: Two immunologically different isozymes of ascorbate peroxidase from spinach leaf
 A:Reference number: S15878; MUID:91378325
 A:Accession: S15878
 A:Molecule type: protein
 A:Residues: 1-3-17 <TAN>
 A>Note: this sequence has been revised in reference S20866
 C:Keywords: chloroplast; oxidoreductase

Query Match 20.6%; Score 21; DB 2; Length 17;
 Best Local Similarity 50.0%; Pred. No. 2.6e+03;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 12 KMPEYKE 19
 | : | | |
 Db 2 KSYPTVHE 9

RESULT 20

S14661
 photosystem I protein psaA - maize (fragment)
 C:Species: Zea mays (maize)
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
 C:Accession: S14661
 R.Kangasjarvi, J.; Gengenbach, B.G.
 submitted to the EMBL Data Library, March 1991
 A:Description: Nucleotide sequence of maize plastid genome BamHI 14 fragment.
 A:Reference number: S14660
 A:Accession: S14661
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-18 <TAN>
 A:Cross-references: EMBL:X58080; NID:g12429; PIDN:CAA41109.1; PID:g12431
 C:Superfamily: Photosystem I P700 apoprotein

Query Match 20.6%; Score 21; DB 2; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2.8e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PEVK 18
 | | | |
 Db 8 PEVK 11

RESULT 21

S02808
 nucleolin - bovine (fragment)
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 18-Jun-1993
 C:Accession: S02808
 R.Sapp, M.; Richter, A.; Weisshart, K.; Calzergues-Ferrer, M.; Amalric, F.; Wallace, M.C.
 Eur. J. Biochem. 179, 541-548, 1989
 A>Title: Characterization of a 48-kDa nucleic-acid-binding fragment of nucleolin.
 A:Reference number: S02808; MUID:89153087
 A:Accession: S02808
 A:Molecule type: protein
 A:Residues: 1-19 <SAP>

Query Match 20.6%; Score 21; DB 2; Length 19;
 Best Local Similarity 45.5%; Pred. No. 2.9e+03;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 IARFKMPEVK 18
 | : | | | |
 Db 3 MAKQKAPPEAK 13

RESULT 22

S57286
 translation elongation factor eEF-1 beta - Sulfolobus solfataricus (fragments)
 C:Species: Sulfolobus solfataricus
 C>Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
 C:Accession: S57286
 R.Arcari, P.; Raimo, G.; Iannicelli, G.; Gallo, M.; Bocchini, V.
 Biochim. Biophys. Acta 1263, 86-88, 1995
 A>Title: The first nucleotide sequence of an archaeal elongation factor 1-beta gene.
 A:Reference number: S57286; MUID:95359209
 A:Accession: S57286
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-20 <ARC>

Query Match 20.6%; Score 21; DB 2; Length 20;
 Best Local Similarity 66.7%; Pred. No. 3.1e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 KMPEY 17
 | : | | |
 Db 8 KVFPEY 13

RESULT 23

S51066
 ribosomal protein S15 - Thermus aquaticus (fragment)
 C:Species: Thermus aquaticus
 C>Date: 15-Jul-1995 #sequence_revision 01-Nov-1996 #text_change 01-Nov-1996
 C:Accession: S51066
 R.Tsibolli, P.; Herfurth, E.; Choli, T.
 Eur. J. Biochem. 226, 169-177, 1994
 A>Title: Purification and characterization of the 30S ribosomal proteins from the bac
 A:Reference number: S51063; MUID:95045586
 A:Accession: S51066
 A:Molecule type: protein
 A:Residues: 1-21 <TSI>
 A>Note: the source is designated as Thermus thermophilus
 C:Keywords: protein biosynthesis; ribosome

Query Match 20.6%; Score 21; DB 2; Length 21;
 Best Local Similarity 44.4%; Pred. No. 3.2e+03;
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 7 MIARFKMP 15
 | : | | | |
 Db 10 VIOEFRRFP 18

RESULT 24

A39108
 lipoprotein TagA precursor - Vibrio cholerae (fragment)
 C:Species: Vibrio cholerae
 C>Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 31-Dec-1993
 C:Accession: A39108
 R.Parsot, C.; Taxman, E.; Mekalanos, J.J.
 Proc. Natl. Acad. Sci. U.S.A. 88, 1641-1645, 1991
 A>Title: ToxR regulates the production of lipoproteins and the expression of serum re
 A:Reference number: A39108; MUID:91156664
 A:Accession: A39108
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-25 <PAR>
 A:Cross-references: GB:M6058
 C:Keywords: lipid binding; lipoprotein

Query Match 20.6%; Score 21; DB 2; Length 25;
 Best Local Similarity 27.3%; Pred. No. 3.9e+03;
 Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 7 MIAFKMPEV 17
| : | : : |
Db 1 MVMVYSLSMKV 11

RESULT 25

S68642
nicotinic acetylcholine receptor-binding protein 4 - black-banded coral snake (fragment)
C:Species: Micrurus nigrocinctus (black-banded coral snake)
C:Date: 23-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 17-Mar-1999
C:Accession: S68642
R:Alape-Giron, A.; Sciles, B.; Schmidt, J.; Giron-Cortes, M.; Thelestam, M.; Joernvall,
FEBS Lett. 380, 29-32, 1996
A:Title: Characterization of multiple nicotinic acetylcholine receptor-binding proteins
A:Reference number: S68639; MUID:96181662
A:Accession: S68642
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <ALA>

Query Match 20.6%; Score 21; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 3.9e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 15 PEVKER 20
| | | : |
Db 14 PEGRDK 19

Search completed: February 5, 2001, 10:49:28
Job time: 741 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:55:56 ; Search time 32.57 Seconds
(without alignments)

19,831 Million cell updates/sec

Title: US-08-981-824-2

Sequence: 102
1 SNMYAMMIAAFKMFPEVKEK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 segs, 32294092 residues

Total number of hits satisfying chosen parameters: 1436

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	28	27.5	15 1 MCRA_METTE	P22948 methanosc
2	25	24.5	14 1 DCMK_PSECF	P19914 pseudomonas
3	23	22.5	10 1 SLAP_BACTG	P49325 bacillus th
4	23	22.5	23 1 UDP_LATCA	P19662 lactobacill
5	23	22.5	24 1 CT31_LITCI	P81851 litorea cit
6	23	22.5	13 1 GRP_SCYCA	P09472 scyllorhinu
7	21	20.6	25 1 PEDT_HYDAT	P80578 hydra atten
8	21	20.6	22 1 CR32_LITCE	P56339 litorea cae
9	21	20.6	22 1 CR33_LITCE	P56340 litorea cae
10	21	20.6	24 1 AMMA_BACTR	P37356 bacillus th
11	20	19.6	15 1 MILT_ONCKE	P81037 oncorhynch
12	20	19.6	18 1 HEMH_THETS	P80155 theromyzon
13	20	19.6	19 1 RL23_HALCU	P05975 halobacteri
14	20	19.6	18 1 LPGE_ECOLI	P33336 escherichia
15	20	19.6	19 1 PSBN_SYNVU	P12313 synchococc
16	20	19.6	20 1 AMP_FUSNU	P81207 fusobacteri
17	20	19.6	24 1 PEPI_ASPNG	P55749 aspergillus
18	20	19.6	25 1 Y15_BPT3	P20835 bacterioph
19	19	18.6	9 1 BUK_CLOPA	P81337 clostridium
20	19	18.6	15 1 UBIL_MONDO	P50103 monodelphis
21	19	18.6	19 1 DHAB_COMTE	P80704 commomops t
22	19	18.6	19 1 NS2_MYCTU	P81136 mycobacteri
23	19	18.6	24 1 CH60_ACTICA	P81874 achinetobact
24	19	18.6	24 1 CLPP_HORVU	P48883 hordeum vul
25	19	18.6	25 1 UBIL_BOVIN	P23356 bos taurus
26	18	17.6	8 1 AL18_CARMA	P81821 carclinus ma
27	18	17.6	10 1 URE3_MORMO	P17339 morangella
28	18	17.6	15 1 UC08_MAIZE	P80614 zea mays (m
29	18	17.6	18 1 ALI2_CYDPO	P82153 cydia pomon
30	18	17.6	20 1 CAOS_RAT	P19633 rattus norv
31	18	17.6	20 1 HG11_FASHE	P80577 fasciola he
32	18	17.6	21 1 PEDB_HYDAT	P80577 hydra atten
33	18	17.6	22 1 AOFI_MOUSE	P06433 mus musculu

34	18	17.6	22 1	CR31_LITSP	P56238 litorea spl
35	18	17.6	22 1	CR34_LITCE	P56241 litorea cae
36	18	17.6	24 1	CPRC_CANFA	P28490 canis famli
37	18	17.6	24 1	IRBP_SHEEP	P12663 ovis aries
38	18	17.6	25 1	FLB1_TREHY	P80158 treponema h
39	18	17.6	25 1	IRBP_PIG	P12662 sus scrofa
40	18	17.6	25 1	PK12_SOLRU	P24744 solanum tub
41	17.5	23 1	COXJ_ONCMY	P80333 oncorhynch	
42	17	16.7	8 1	AL12_CARMA	P81819 carclinus ma
43	17	16.7	9 1	PGLR_DIAAB	P81179 diatrepes a
44	17	16.7	9 1	SAP_STOVA	P24047 stomopneute
45	17	16.7	12 1	PA2B_VIPBO	P31859 vipera beru

ALIGNMENTS

RESULT 1	MCRA_METTE	STANDARD:	PRT:	15 AA.
ID	MCRA_METTE			
AC	P22948;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	METHYL-COENZYME M METHYLRREDUCTASE ALPHA SUBUNIT (EC 1.8.-.-)			
DE	(FRAGMENT).			
OS	Methanosarcina thermophila.			
OC	Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;			
OC	Methanosarcina.			
RN	[1]			
RP	SEQUENCE.			
RC	STRAIN-DSM 1825 / TM-1;			
RX	MEDLINE-91193204; PubMed-2013570;			
RA	Jablonski P.E., Ferry J.G.;			
RT	"Purification and properties of methyl coenzyme M methylreductase from acetate-grown Methanosarcina thermophila.";			
RT	J. Bacteriol. 173:2481-2487(1991).			
CC	-1- FUNCTION: THIS ENZYME COMPLEX CATALYZES THE FINAL STEP IN			
CC	METHANOGENESIS, WHICH IS THE TERMINAL STEP OF ANAEROBIC			
CC	DEGRADATION OF BIOMASS.			
CC	-1- CATALYTIC ACTIVITY: REDUCTION OF METHYL-COENZYME M (2-(METHYLTHIO)			
CC	ETHANESULFONIC ACID) WITH 7-MERCAPTOHEPTANOYLTHREONINE PHOSPHATE			
CC	TO METHANE AND AN HETERODISULFIDE (CH(3)-S-COM + H-S-HTP - CH(4) +			
CC	COM-S-S-HTP).			
CC	-1- COFACTOR: THE ENZYME COMPLEX BINDS TIGHTLY (BUT NOT COVALENTLY)			
CC	TO ONE MOLECULE OF COENZYME F430. F430 IS A YELLOW NICKEL			
CC	PORPHINOID.			
CC	-1- SUBUNIT: TRIMER OF AN ALPHA, A BETA, AND A GAMMA SUBUNITS.			
CC	-1- MISCELLANEOUS: THE OPTIMAL TEMPERATURE FOR ENZYME ACTIVITY IS 60			
CC	DEGREES CELSIUS.			
KW	MISCELLANEOUS: REDUCED FERREDOXIN COULD REDUCTIVELY REACTIVATE THE			
KW	ENZYME.			
FT	Methanogenesis; Oxidoreductase.			
FT	NON TER 15 15			
SO	SEQUENCE 15 AA; 1686 MM; DSD59892FCA5F63C CRC64;			
Query Match 27.5%; Score 28; DB 1; Length 15;				
Best Local Similarity 50.0%; Pred. No. 65;				
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;				
QY	5 AMMIARFKMFPEVK 18			
DB	1: 1:11 111			
2 ADIFAKFTSMVEVK 15				
RESULT 2				
ID	DCMK_PSECF	STANDARD:	PRT:	14 AA.
AC	P19914;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	01-DEC-1992 (Rel. 24, Last annotation update)			

```

DE CARBON MONOXIDE DEHYDROGENASE MEDIUM CHAIN (EC 1.2.99.2) (FRAGMENT).
OS Pseudomonas carboxydoflava.
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
CC Hydrogenophaga.
RN [1]
RP SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
  carboxydophilic bacteria."
RL Arch. Microbiol. 152:335-341(1989).
CC -1- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
  ACCEPTOR.
CC -1- COFACTOR: MOLYBDENUM.
CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
  SMALL.
CC PIR: P10142; P10142.
KM Oxidoreductase; Molybdenum.
FT NON_TER 14
SQ SEQUENCE 14 AA; 1756 MW; 65583C6D1FB7C25B CRC64;

Query Match 24.5%; Score 25; DB 1; Length 14;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 MIAERK 12
   |||
   1 MIAERF 7

RESULT 3
SLAP_BACTG
ID SLAP_BACTG STANDARD; PRT; 10 AA.
AC P49323;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE S-LAYER PROTEIN (SURFACE LAYER PROTEIN) (FRAGMENT).
OS Bacillus thuringiensis (subsp. galleriae).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
  Bacillus/staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE.
RC STRAIN=NRRL 4045;
RX MEDLINE=90078111; PubMed=2592346;
RA Luckevich M.D., Beveridge T.J.;
RT "Characterization of a dynamic S layer on Bacillus thuringiensis.";
RL J. Bacteriol. 171:6656-6667(1989).
CC -1- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
  OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC -1- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
  CELL WALL; S-LAYER.
KM NON_TER 10
SQ SEQUENCE 10 AA; 1080 MW; 57AECACAB769D1A3 CRC64;

Query Match 22.5%; Score 23; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 KMPEV 17
   |||
   3 KTFPV 8

RESULT 4
UDP_LACCA
ID UDP_LACCA STANDARD; PRT; 23 AA.
AC P19662;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)

```

```

DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE URIDINE PHOSPHORYLASE (EC 2.4.2.3) (UDRPASE) (FRAGMENT).
GN UDP.
OS Lactobacillus casei.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
  Lactobacillus.
RN [1]
RP SEQUENCE.
RX STRAIN=ATCC 7469; PubMed=2119230;
RA Avraham Y., Grossowicz N., Yashpe J.;
RT "Purification and characterization of uridine and thymidine
  phosphorylase from Lactobacillus casei."
RL Biochim. Biophys. Acta 1040:287-293(1990).
CC -1- FUNCTION: THE ENZYMES WHICH CATALYZE THE REVERSIBLE PHOSPHORYLOSIS
  OF PYRIMIDINE NUCLEOSIDES ARE INVOLVED IN THE DEGRADATION OF THESE
  COMPOUNDS AND IN THEIR UTILIZATION AS CARBON AND ENERGY SOURCES,
  OR IN THE RESCUE OF PYRIMIDINE BASES FOR NUCLEOTIDE SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: URIDINE + PHOSPHATE = URACIL + ALPHA-D-RIBOSE
  1-PHOSPHATE.
CC -1- PATHWAY: NUCLEOTIDE AND DEOXYRIBONUCLEOTIDE CATABOLISM.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE PNP/UDP FAMILY 1 OF PHOSPHORYLASES.
DR PIR: S11383; S11383.
DR INTERPRO: IPR000845;
DE PROSITE, PS01232; PNP_UDP_1; PARTIAL.
KM Transferase; Glycosyltransferase
FT UNSURE 1
FT UNSURE 7
FT NON_TER 23
SQ SEQUENCE 23 AA; 2630 MW; 918B2E2F32F35A17 CRC64;

Query Match 22.5%; Score 23; DB 1; Length 23;
Best Local Similarity 57.4%; Pred. No. 7.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 RFRKPE 16
   |||
   1 RLVKVE 7

```

```

RESULT 5
CT31_LITCI
ID CT31_LITCI STANDARD; PRT; 24 AA.
AC P81851; P81852; P81853;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CITROPIN 3.1.2 (CONTAINS: CITROPIN 3.1.1; CITROPIN 3.1).
OS Litoria citropa (Australian blue mountains tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
  Litoria.
RN [1]
RP SEQUENCE.
RC TISSUE=SKIN;
RX MEDLINE=99435977; PubMed=10504394;
RA Wegener K.L., Wabnitz P.A., Carver J.A., Bowle J.H., Chia B.C.S.,
  Wallace J.C., Tyler M.J.;
RT "Wallace J.C., Tyler M.J.;
  "Halt defense peptides from the skin glands of the Australian blue
  RT mountain tree frog Litoria citropa. Solution structure of the
  RT antibacterial peptide citropin 1.1."
RL Eur. J. Biochem. 265:627-637(1999).
CC -1- TISSUE SPECIFICITY: DORSAL AND SUBVENTRAL SKIN GLANDS.
FT PEPTIDE 1 24 CITROPIN 3.1.2.
FT PEPTIDE 1 23 CITROPIN 3.1.1.
FT PEPTIDE 1 22 CITROPIN 3.1.
SQ SEQUENCE 24 AA; 2614 MW; C9001E295BD0E15D CRC64;

Query Match 22.5%; Score 23; DB 1; Length 24;
Best Local Similarity 50.0%; Pred. No. 7.7e+02;

```

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 13 MEPEVEK 20

DB 2 LEQVIREK 9

RESULT 6

GRP_SCYCA STANDARD; PRT; 25 AA.
ID GRP_SCYCA
AC P09472;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE GASTRIN-RELEASING PEPTIDE (GRP).
OS Scyliorhinus canicula (Spotted dogfish). (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeidae; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
RN [1]
RP SEQUENCE.
RX MEDLINE-88137922; PubMed-3436516;
RA Conlon J.M., Henderson I.W., Thim L.;
RT "Gastrin-releasing peptide from the intestine of the elasmobranch
fish, Scyliorhinus canicula (common dogfish).";
RL Gen. Comp. Endocrinol. 68:415-420(1987).
CC -1- FUNCTION: GRP STIMULATES GASTRIN RELEASE AS WELL AS OTHER
CC GASTROINTESTINAL HORMONES.
CC -1- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
CC FAMILY.
DR PIR: S06263; S06263.
DR INTERPRO: IPRO00874; -.
DR PRAM: P02044; Bombesin; 1.
DR PROSITE: PS00257; BOMBESIN; 1.
KW Bombesin family; Amidation.
FT MOD.RES 25 25
SO SEQUENCE 25 AA; 2781 MW; B735F911B89007F8 CRC64;

Query Match 22.5%; Score 23; DB 1; Length 25;

Best Local Similarity 100.0%; Pred. No. 8e+02; Mismatches 0; Indels 0; Gaps 0;

QY 12 KMFP 15

DB 11 KMFP 14

RESULT 7

PEDI_HYDAT STANDARD; PRT; 13 AA.
ID PEDI_HYDAT
AC P80578;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PEDI.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
RN [1]
RP SEQUENCE.
RX MEDLINE-96232307; PubMed-8674432;
RA Hofmeister S.A.H.;
RT "Isolation and characterization of two new morphogenetically active
peptides from Hydra vulgaris.";
RL Development 122:1941-1948(1996).
CC -1- FUNCTION: MORPHOGENETICALLY ACTIVE PEPTIDE. ACTIVE IN FOOT
CC DEVELOPMENT.
KW Morphogen.
SO SEQUENCE 13 AA; 1512 MW; 6F5266E3F37632CB CRC64;

Query Match 20.6%; Score 21; DB 1; Length 13;

Best Local Similarity 40.0%; Pred. No. 9.1e+02; Mismatches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 REFMEPEVEK 19

DB 4 RPEVLPDVE 13

RESULT 8

CR32_LITCE STANDARD; PRT; 22 AA.
ID CR32_LITCE
AC P56239;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CAERIN 3.2.
OS Litoria caerulea.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-PAROTOID GLAND;
RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins from
Litoria caerulea.";
RL J. Chem. Res. 138:910-936(1993).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL
CC CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC GLANDS.
CC -1- MASS SPECTROMETRY: MW-2397; METHOD-FAB.
KW Antibiotic; Amphibian skin; Amidation.
FT MOD.RES 22 22
SO SEQUENCE 22 AA; 2400 MW; 1D440B3829D4367C CRC64;

Query Match 20.6%; Score 21; DB 1; Length 22;

Best Local Similarity 37.5%; Pred. No. 1.5e+03; Mismatches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 13 MEPEVEK 20

DB 2 LMEKIREK 9

RESULT 9

CR33_LITCE STANDARD; PRT; 22 AA.
ID CR33_LITCE
AC P56240;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CAERIN 3.3.
OS Litoria caerulea.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-PAROTOID GLAND;
RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins from
Litoria caerulea.";
RL J. Chem. Res. 138:910-936(1993).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL
CC CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC GLANDS.
CC -1- MASS SPECTROMETRY: MW-2424; METHOD-FAB.
KW Antibiotic; Amphibian skin; Amidation.

FT MOD_RES 22 22 AMIDATION.
SQ SEQUENCE 22 AA: 2427 MW: 1D440B2200D4367C CRC64;

Query Match 20.6%; Score 21; DB 1; Length 22;
Best Local Similarity 37.5%; Pred. No. 1.5e+03;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 13 MFPEVKE 20
: : : : :
Db 2 LMEKIKER 9

RESULT 10
ID AMAA_BACTR STANDARD; PRT; 24 AA.
AC P37356;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14) (L-AMINOACYLASE)
OS (FRAGMENT).
OC Bacillus thermoglucosidasius.
OC Bacillus firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RA Chu H.-Y., Tanizawa K., Tanaka H., Soda K.;
RT "Thermostable aminoacylase from Bacillus thermoglucosidasius.
RT Purification and characterization.";
RL Agric. Biol. Chem. 51:2793-2800(1987).
CC -1- CATALYTIC ACTIVITY: A N-ACYL-L-AMINO ACID + H(2)O -> A FATTY ACID
CC ANION + A L-AMINO ACID.
CC -1- COFACTOR: ACTIVATED BY COBALT (BY SIMILARITY).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M40, ALSO KNOWN AS THE
CC AMA/HIPO/HYVC FAMILY OF HYDROLASES.
KM Hydrolyase; Cobalt.
FT NON_TER 24
SQ SEQUENCE 24 AA: 2846 MW: BC954EAD2B0EC64A CRC64;

Query Match 20.6%; Score 21; DB 1; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 12 KMPEVKE 19
: : : : :
Db 8 RLVDVEKE 15

RESULT 11
ID MILT_ONCKE STANDARD; PRT; 15 AA.
AC P81037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MULTIPAIN (EC 3.4.22.-) (FRAGMENT).
OS Oncorhynchus keta (Chum salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RA MEDLINE=97397031; PubMed=9253183;
RA Kawabata C., Ichishima E.;
RT "Miltipain, new cysteine proteinase from the milt of chum salmon,
RT Oncorhynchus keta";
RL Comp. Biochem. Physiol. 117B:445-452(1997).
CC -1- FUNCTION: CYSTEINE PROTEINASE THAT HYDROLYSES BASIC PROTEINS.
CC HYDROLYSE BASIC PROTEINS SUCH AS HISTONE, SALMINE AND CLUPAINE BUT
CC NOT MILK CASEIN.

CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE WITH BASIC RESIDUES AT
CC P2 AND P1.
CC MEROPS; C01.093; -.
KM Hydrolyase.
DT NON_TER 15
SQ SEQUENCE 15 AA: 1730 MW: 766B7771C0F888E7 CRC64;

Query Match 19.6%; Score 20; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNMYAMMI 8
: : : : :
Db 3 SFLYAENV 10

RESULT 12
ID HEMH_THETS STANDARD; PRT; 18 AA.
AC P80155;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE OVOHEMERETHRIN (YP14) (FRAGMENT).
OS Theromyzon tessulatum (leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea;
OC Rhynchobdellida; Glossiphoniidae; Theromyzon.
RN [1]
RP SEQUENCE.
RC TISSUE=OOCYTE;
RX MEDLINE=93049299; PubMed=1425663;
RA Bert J.-L., Britel M., Sautiere P., Malecha J.;
RT "Ovohemerethrin, a major 14-kDa yolk protein distinct from
RT vitellogenin in leech.";
RL Eur. J. Biochem. 209:563-569(1992).
CC -1- FUNCTION: MAJOR YOLK PROTEIN. THIS IRON PROTEIN MAY PLAY A ROLE
CC IN THE DETOXIFICATION OF FREE IRON AFTER A BLOOD MEAL.
CC -1- SIMILARITY: BELONGS TO THE HEMERYTHRIN FAMILY.
CC PIR; S29264; S29264.
DR HSSP; P02247; ZMR.
DR INTERPRO; IPR002063; -.
DR PROSITE; PS00550; HEMERYTHRINS; PARTIAL.
KM Oxygen transport; Metal-binding; Iron; Yolk.
FT NON_TER 18
SQ SEQUENCE 18 AA: 2368 MW: 33397EE587C81F1 CRC64;

Query Match 19.6%; Score 20; DB 1; Length 18;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKMPPE 16
: : : : :
Db 13 FKVFYE 18

RESULT 13
ID RL23_HALCU STANDARD; PRT; 18 AA.
AC P05975;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L23P (HL31) (FRAGMENT).
GN RPL23P.
OS Halobacterium cutirubrum.
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
RN [1]
RP SEQUENCE.
RX MEDLINE=84282108; PubMed=6467081;
RA Matheson A.T., Yaguchi M., Christensen P., Rollin C.F., Hasnain S.;

RT "Purification, properties, and N-terminal amino acid sequence of
RT certain 50S ribosomal subunit proteins from the archaeobacterium
RT Halobacterium cutirubrum.";
RL Can. J. Biochem. Cell Biol. 62:426-433(1984).
CC -1- FUNCTION: BINDS TO A SPECIFIC REGION ON THE 23S RNA
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
DR PIR: S08559; S08559.
DR INTERPRO: IPR001014;
DR PROSITE: PS00050; RIBOSOMAL_L23; PARTIAL.
KW Ribosomal protein; RNA-binding.
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 2041 MW; 3E61DC53F8B4DD4C CRC64;

Query Match 19.6%; Score 20; DB 1; Length 18;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 14 EPEVKR 20
: | | | |
DB 6 YPLVTEK 12

RESULT 14
LPG_ECOLI
ID LPG_ECOLI STANDARD; PRT; 19 AA.
AC P33236;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GEF LEADER PEPTIDE.
GN GEF.

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.

RP SEQUENCE FROM N.A.
RA MEDLINE-92048481; PubMed-1943701;
RA Poulsen L.R., Refn A., Molin S., Andersson P.;
RT "The gef gene from Escherichia coli is regulated at the level of
RT translation";
RL Mol. Microbiol. 5:1639-1648(1991).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: AE000112; AAC73129.1; ALT_TERM.
DR PIR: S16473; S16473.
DR ECOGENE: EGI2074; gefL.
KW Leader peptide.
FT
SQ SEQUENCE 19 AA; 2259 MW; 19B3EDF371EB0BB CRC64;

Query Match 19.6%; Score 20; DB 1; Length 19;
Best Local Similarity 80.0%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 16 EYVKR 20
: | | | |
DB 13 KYVKR 17

RESULT 15
PSBN_SYNVU
ID PSBN_SYNVU STANDARD; PRT; 19 AA.
AC P12313;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PHOTOSYSTEM II REACTION CENTER N PROTEIN (FRAGMENT).
GN PSBN.
OS Synechococcus vulcanus.
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.

RP SEQUENCE.
RX MEDLINE-89338735; PubMed-2503398;
RA Ikeuchi M., Koike H., Inoue Y.;
RT "N-terminal sequencing of low-molecular-mass components in
RT cyanobacterial photosystem II core complex. Two components correspond
RT to unidentified open reading frames of plant chloroplast DNA.";
RL FEBS Lett. 253:178-182(1989).

CC -1- FUNCTION: NOT KNOWN.
CC -1- SIMILARITY: BELONGS TO THE PSBN FAMILY.
DR PIR: S05216; S05216.
KW Photosystem II; Transmembrane.
FT MOD_RES 1 1 BLOCKED.
FT TRANSMEM 6 >19 POTENTIAL.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2217 MW; A97C99B523106D14 CRC64;

Query Match 19.6%; Score 20; DB 1; Length 19;
Best Local Similarity 44.4%; Pred. No. 2e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 MYAMTARF 11
: | | | |
DB 9 IFAXITLALF 17

RESULT 16
AMP_FUSNU
ID AMP_FUSNU STANDARD; PRT; 20 AA.
AC P81207;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE AMINOPEPTIDASE (EC 3.4.11.-) (AP) (FRAGMENT).
OS Fusobacterium nucleatum.
OC Bacteria; Fusobacteria; Fusobacterium.

RP SEQUENCE.
RC STRAIN-ATCC 10953;
RX MEDLINE-98361026; PubMed-9695913;
RA Rogers A.H., Gunadi A., Gully N.J., Zilm P.S.;
RT "An aminopeptidase nutritionally important to Fusobacterium
RT nucleatum";
RL Microbiology 144:1807-1813(1998).
CC -1- FUNCTION: CLEAVES A WIDE RANGE OF DIPEPTIDES AND TRIPEPTIDES.
CC -1- COFACTOR: COBALT.
CC -1- SUBCELLULAR LOCATION: CELL ENVELOPE-ASSOCIATED.
KW Hydrolase; Aminopeptidase; Metalloprotease; Cobalt.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2585 MW; 97095B948262C71B CRC64;

Query Match 19.6%; Score 20; DB 1; Length 20;
Best Local Similarity 33.3%; Pred. No. 2.1e+03;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

OY      12 KMPEVKEK 20
DB      4 KXYVLEKER 12

RESULT 17
PEPL_ASPNG STANDARD: PRT: 24 AA.
AC P55749;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE CARBOXYPEPTIDASE I (EC 3.4.16.-) (CPD-1) (FRAGMENT).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae;
OC anamorphic trichocomaceae; Aspergillus.
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE-92344370; PubMed-1637154;
RA Dal Degam F., Ribadeau-Dumas B., Bredem K.;
RT Purification and characterization of two serine carboxypeptidases
RT from Aspergillus niger and their use in C-terminal sequencing of
RT proteins and peptide synthesis.
RL Appl. Environ. Microbiol. 58:2144-2152(1992).
CC -1- FUNCTION: REMOVES ACIDIC, NEUTRAL AND BASIC AMINO ACIDS AS WELL AS
CC PROLINE FROM THE C-TERMINAL POSITION. DIGESTS PREFERENTIALLY
CC PEPTIDES CONTAINING A HYDROPHOBIC RESIDUE IN P1 POSITION, AS WELL
CC AS ARGININE, LEUCINE OR PHENYLALANINE IN P1 POSITION OF ESTER
CC SUBSTRATE. OPTIMAL PH IS 4; ENZYME IS UNSTABLE ABOVE PH 8.
CC CATALYZES ALSO PEPTIDE SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF A C-TERMINAL AMINO ACID WITH A
CC BROAD SPECIFICITY.
CC -1- ENZYME REGULATION: INHIBITED BY DFP.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- PTM: CONTAINS BOTH N- AND O-LINKED SUGAR CHAINS. THE N-LINKED
CC OLIGOSACCHARIDES ARE UNIQUE STRUCTURES OF MAN(10)GLCNAc(2) AND
CC MAN(11)GLCNAc(2). DEGLYCOSYLATION DOES NOT AFFECT THE CATALYTIC
CC ACTIVITY. PH OR THERMAL STABILITY, OR RESISTANCE TO PROTEOLYSIS OF
CC THE ENZYME.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10; ALSO KNOWN AS THE
CC SERINE CARBOXYPEPTIDASE FAMILY.
DR INTERPRO: IPR001563;
DR PROSITE: PS00131; CARBOXYPEPT_SER_SER; PARTIAL.
DR PROSITE: PS00560; CARBOXYPEPT_SER_HIS; PARTIAL.
KW Hydrolase; Carboxypeptidase; Glycoprotein.
FT UNSURE 2
FT UNSURE 11 11 PROBABLE.
FT CARBOHYD 2 2
FT CARBOHYD 11 11 N-LINKED (GLCNAc. . .) (PROBABLE).
FT NON_TER 24 24
SO SEQUENCE 24 AA; 2623 MW; B7F183DEF6DBA6C2 CRC64;

Query Match 19.6%; Score 20; DB 1; Length 24;
Best Local Similarity 29.4%; Pred. No. 2.5e+03;
Matches 5; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

OY      1 SNMTAMTAREKPEV 17
DB      2 TNKTAFLVNGTSIPEV 18

RESULT 18
ID Y15_BP13 STANDARD: PRT: 25 AA.
AC P20835;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 01-FEB-1991 (Rel. 17, Last annotation update)
DE HYPOTHETICAL GENE 1.5 PROTEIN.
GN 1.5.
OS Bacteriophage T3.

```

```

OC      17-1-like phages.
OC      17-1-like phages.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LURIA;
RX MEDLINE-87226207; PubMed-3586029;
RA Schmitt M.P., Beck P.J., Kearney C.A., Spence J.L., Digiovanni D.,
RA Condreay J.P., Molinoux I.J.;
RT "Sequence of a conditionally essential region of bacteriophage T3,
RT including the primary origin of DNA replication."
RL J. Mol. Biol. 193:479-495(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17255; CA35126.1;
DR EMBL; X05031; CAA28701.1;
DR PIR; S09540; S09540.
KW Hypothetical protein.
SO SEQUENCE 25 AA; 2821 MW; 92B709153A2950A CRC64;

Query Match 19.6%; Score 20; DB 1; Length 25;
Best Local Similarity 27.3%; Pred. No. 2.6e+03;
Matches 3; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY      6 MMIRFKPE 16
DB      14 LMVADNINWD 24

RESULT 19
ID BURK_CLOPA STANDARD: PRT: 9 AA.
AC P8137;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE BUTYRATE KINASE (EC 2.7.2.7) (BK) (CP 38) (FRAGMENT).
GN BURK.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
RN [1]
RP SEQUENCE.
RC STRAIN-W5;
RX MEDLINE-98291870; PubMed-9629918;
RA Flengrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5."
RL Electrophoresis 19:802-806(1998).
CC -1- FUNCTION: CATALYZES THE CONVERSION OF BUTYRYL-COA THROUGH BUTYRYL
CC PHOSPHATE TO BUTYRATE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + 2-BUTANOATE -> ADP + BUTANOYL PHOSPHATE.
CC -1- PATHWAY: INVOLVED IN THE ACIDOGENIC PHASE OF FERMENTATION.
CC -1- SIMILARITY: BELONGS TO THE ACETOKINASE FAMILY.
DR INTERPRO: IPR000890;
DR PROSITE: PS01075; ACETATE_KINASE_1; PARTIAL.
DR PROSITE: PS01076; ACETATE_KINASE_2; PARTIAL.
KW Transferase; Kinase.
FT NON_TER 9 9
SO SEQUENCE 9 AA; 1104 MW; 055D40505727232B CRC64;

Query Match 18.6%; Score 19; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 8.9e+04;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```



```

RT Acinetobacter calcoaceticus";
RL Electrophoresis 20:781-789(1999).
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- INDUCTION: BY HEAT SHOCK AND ETHANOL.
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
DR INTERPRO: IPR001844; -.
DR PROSITE: PS00296; CHAPERONIN_CPN60; PARTIAL.
KM Chaperone; ATP-binding; Heat shock.
FT NON_TER 24
SQ SEQUENCE 24 AA: 2483 MW: 83029B180D735688 CRC64;

Query Match 18.6%; Score 19; DB 1; Length 24;
Best Local Similarity 80.0%; Pred. No. 3.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 AMMTA 9
DB 13 SMTA 17

RESULT 24
CLPP_HORVU STANDARD; PRT; 24 AA.
AC P48883;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROBABLE CLPP-LIKE PROTEASE (EC 3.4.21.92) (ENDOPEPTIDASE CLP)
DE (FRAGMENT).
GN CLPP.
OS Hordeum vulgare (Barley).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. HAISA;
RX MEDLINE=96197402; Pubmed=8616228;
RA Huebschmann T., Hess W., Boerner T.;
RT "Impaired splicing of the rps12 transcript in ribosome-deficient
RT plastids.";
RL Plant Mol. Biol. 30:109-123(1996).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS TO SMALL PEPTIDES IN
CC THE PRESENCE OF ATP AND MAGNESIUM. ALPHA-CASEIN IS THE USUAL TEST
CC SUBSTRATE. IN THE ABSENCE OF ATP, ONLY OLIGOPEPTIDES SHORTER THAN
CC FIVE RESIDUES ARE CLEAVED (SUCH AS SUCCINYL-LEU-TYR-|-NHMEC; AND
CC LEU-TYR-|-LEU-|-TYR-TRP, IN WHICH THE CLEAVAGE OF THE -TYR-|-LEU-
CC AND -TYR-|-TRP- BOND ALSO OCCURS).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S14; ALSO KNOWN AS CLPP
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X89562; CA61738.1; -.
CC MENDEL: 2297; HORVU:clpp.1.
CC INTERPRO: IPR001907; -.
CC PROSITE: PS00381; CLP_PROTEASE_SER; PARTIAL.
CC PROSITE: PS00382; CLP_PROTEASE_HIS; PARTIAL.
CC HYDROLASE; Serine protease; Chloroplast.
FT NON_TER 1
SQ SEQUENCE 24 AA: 2941 MW: 715E01FE3E073CE CRC64;

```

```

Query Match 18.6%; Score 19; DB 1; Length 24;
Best Local Similarity 75.0%; Pred. No. 3.7e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 FPEV 17
DB 17 FPEM 20

RESULT 25
UBLI_BOVIN STANDARD; PRT; 25 AA.
AC P23356;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE ISOZYME L1 (EC 3.1.2.15) (UCH-
DE L1) (UBIQUITIN THIOLESTERASE L1) (NEURON CYTOPLASMIC PROTEIN 9.5)
DE (PGP 9.5) (FRAGMENT).
GN UCHL1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE.
RC TISSUE=BRAIN;
RX MEDLINE=92008646; Pubmed=1833240;
RA Giampanco I., Bianchi R., Caccarelli P., Pula G., Sorci G.,
RA Antonoli S., Bocchini V., Donato R.;
RT "Neuron-specific" protein gene product 9.5 (PGP 9.5) is also
RT expressed in glioma cell lines and its expression depends on cellular
RT growth state.";
RL FEBS Lett. 290:131-134(1991).
CC -1- FUNCTION: UBIQUITIN-PROTEIN HYDROLASE IS INVOLVED BOTH IN THE
CC PROCESSING OF UBIQUITIN PRECURSORS AND OF UBIQUINATED PROTEINS.
CC THIS ENZYME IS A THIOLESTERASE THAT RECOGNIZE AND HYDROLYZE
CC A PEPTIDE BOND AT THE C-TERMINAL GLYCINE OF UBIQUITIN.
CC -1- CATALYTIC ACTIVITY: UBIQUITIN C-TERMINAL THIOLESTER + H(2)O -
CC UBIQUITIN + A THIOLE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: NEURONS AND CELLS OF THE DIFFUSE
CC NEUROENDOCRINE SYSTEM AND THEIR TUMORS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C12; ALSO KNOWN AS
CC FAMILY 1 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.
CC PIR: S17561; S17561.
DR MEROPS: C12.001; -.
DR INTERPRO: IPR001578; -.
DR PROSITE: PS00140; UCH_1; PARTIAL.
KM Ubiquitin conjugation; Hydrolase; Thiol protease; Multigene family.
FT NON_TER 25
SQ SEQUENCE 25 AA: 2812 MW: 26BB5AD0A754D55 CRC64;

Query Match 18.6%; Score 19; DB 1; Length 25;
Best Local Similarity 26.7%; Pred. No. 3.8e+03;
Matches 4; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 6 MIAFEPKEVEKER 20
DB 1 MQLKPMIDEMLKNK 15

```

Search completed: February 5, 2001, 10:55:58
Job time: 498 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:55:20 ; Search time 93.77 Seconds
(without alignments)
24.999 Million cell updates/sec

Title: US-08-981-824-2

Perfect score: 102
Sequence: 1 SNRYAMAMARFKMEPEVKEK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.15:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_undefined:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	29.4	23	6	Q9R314
2	26	25.5	17	6	Q9TRH5
3	26	25.5	19	11	Q9J1A2
4	26	25.5	22	11	Q9QW14
5	24	23.5	11	11	Q9QXN6
6	24	23.5	18	2	Q9RGRI
7	24	23.5	21	8	Q35556
8	23	22.5	14	11	Q9J1U5
9	23	22.5	17	12	Q85673
10	23	22.5	17	12	Q00340
11	23	22.5	18	2	Q9R4N5
12	23	22.5	20	2	Q9R4W0
13	23	22.5	22	2	Q05279
14	22	22.5	22	2	Q9S551
15	22	21.6	12	11	Q62966
16	22	21.6	15	1	Q9UWM1
17	22	21.6	15	8	Q35921
18	22	21.6	16	4	Q9UMJ9
19	22	21.6	20	2	Q9RLP2

20	22	21.6	20	4	Q9UK71
21	22	21.6	20	6	Q29341
22	22	21.6	20	6	Q9TRC0
23	22	21.6	21	6	Q9XSL5
24	22	21.6	24	2	Q9ZBO5
25	22	21.6	24	2	Q9XAK2
26	22	21.6	24	2	Q9R4T3
27	22	21.6	24	6	Q97793
28	22	21.6	25	2	Q9R5D7
29	22	21.6	25	5	Q76530
30	22	21.6	25	11	Q9JK04
31	22	21.6	25	12	Q9TER9
32	21.5	21.1	20	6	Q9TQX5
33	21.5	21.1	25	6	Q77646
34	21	20.6	13	2	Q9RF24
35	21	20.6	15	2	Q9RG03
36	21	20.6	15	6	Q9TR14
37	21	20.6	17	2	Q52748
38	21	20.6	17	8	Q9XQ89
39	21	20.6	18	8	Q19969
40	21	20.6	18	8	Q19979
41	21	20.6	18	8	Q19971
42	21	20.6	18	8	Q19975
43	21	20.6	18	8	Q19973
44	21	20.6	19	8	Q19977
45	21	20.6	19	2	Q9RF27

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	23 AA.
Q9R314	Q9R314	Q9R314		
AC	Q9R314	Q9R314		
DT	01-MAY-2000 (TREMREL. 13, Created)			
DT	01-MAY-2000 (TREMREL. 13, Last sequence update)			
DT	01-JUN-2000 (TREMREL. 14, Last annotation update)			
DE	DNA ADENINE METHYLASE HOMOLOG (FRAGMENT).			
GN	M. HPI.			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria: Proteobacteria; epsilon subdivision; Helicobacter group;			
CC	Helicobacter.			
OX	NCBI_TaxID=210;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-219, AND 213;			
RA	Raudonkluene A., Berg D.E.;			
RT	"IceA2 segment from H. pylori (Alaska strain 219)."			
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF008929; AAC64502.1;			
DR	EMBL; AF008928; AAC64500.1;			
KW	Methyltransferase.			
FT	NON-TER			
SO	SEQUENCE	23 AA: 2635 MW: 6D6CC63737422B3 CRC64;		

Query Match 29.4%; Score 30; DB 2; Length 23;
Best Local Similarity 36.4%; Pred. No. 2.5e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY	9	ARFKMEPEVKE 19
DB	6	SKYRLPIRKE 16
Q9TRH5	2	
AC	Q9TRH5	PRELIMINARY: PRT: 17 AA.
DT	01-MAY-2000 (TREMREL. 13, Created)	
DT	01-MAY-2000 (TREMREL. 13, Last sequence update)	
DT	01-JUN-2000 (TREMREL. 14, Last annotation update)	

DE ALPHA-SI-CASEIN HOMOLOG (FRAGMENT).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN (1)
 RP SEQUENCE.
 RX MEDLINE=93231344; PubMed=1299613;
 RA Neuhoom B., Giuffrida M.G., Cont A.;
 RT "Isolation of a new ligand-carrying casein fragment from bovine
 RT mammary gland microsomes";
 RT FEBS Lett. 305:189-191(1992).
 SQ SEQUENCE 17 AA; 1950 MW; 8890EF0685A0537F CRC64;

Query Match 25.5%; Score 26; DB 6; Length 17;
 Best Local Similarity 77.8%; Pred. No. 8.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

OY 14 FPEV--KER 20
 ||| ||
 Db 5 FPEVGEK 13

RESULT 3
 ID 09JIA2 PRELIMINARY; PRT; 19 AA.
 AC 09JIA2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE NORPEPHRINE TRANSPORTER (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Padbury J.F., McGonigal B., Tseng Y.;
 RT "Cloning and Sequence Analysis of the Rat Norepinephrine Transporter
 RT Promoter";
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF246668; AAF78041.1; -;
 FT NON_TER 19
 SQ SEQUENCE 19 AA; 2082 MW; 4AA83339008F55CD CRC64;

Query Match 25.5%; Score 26; DB 11; Length 19;
 Best Local Similarity 46.2%; Pred. No. 9.6e+02;
 Matches 6; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

OY 6 MMIAFFKPEVK 18
 ||:| | | | |
 Db 1 MLIAKMK--PQVQ 11

RESULT 4
 ID 09QM14 PRELIMINARY; PRT; 22 AA.
 AC 09QM14;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PROTEIN TYROSINE KINASE JAK1 (FRAGMENT).
 OS Mus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10095;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93232338; PubMed=8514334;

RA Harpur A.G., Zimiecki A., Wilks A.F., Falk K., Rotzschke O.,
 RA Rammensee H.G.;
 RT "A prominent natural H-2 Kd ligand is derived from protein tyrosine
 RT kinase JAK1";
 RL Immunol. Lett. 35:235-237(1993).
 SQ SEQUENCE 22 AA; 2681 MW; D0110BD1FC3C084B CRC64;

Query Match 25.5%; Score 26; DB 11; Length 22;
 Best Local Similarity 57.1%; Pred. No. 1.1e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 11 FKMFPEV 17
 | |||
 Db 6 FSFPEI 12

RESULT 5
 ID 09QXN6 PRELIMINARY; PRT; 11 AA.
 AC 09QXN6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE INTERLEUKIN-2 (FRAGMENT).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Lyons P.A., Amelage N., Lord C.J., Denny P., Hill N.J., Podolin P.L.,
 RA Peterson L.B., Wicker L.S., Todd J.A.;
 RT "A 780 Kb physical map of proximal mouse chromosome 3 encompassing the
 RT type 1 diabetes locus, Id43";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF195955; AAF22748.1; -;
 FT NON_TER 11
 SQ SEQUENCE 11 AA; 1233 MW; AE3BEC9CBB736D6 CRC64;

Query Match 23.5%; Score 24; DB 11; Length 11;
 Best Local Similarity 57.1%; Pred. No. 1.2e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 MYAMMIA 9
 ||:| | |
 Db 1 MYSMOLA 7

RESULT 6
 ID 09RGRI PRELIMINARY; PRT; 18 AA.
 AC 09RGRI;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE ADHESIN (FRAGMENT).
 GN FHAB.
 OS Bordetella parapertussis.
 CC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 OX NCBI_TaxID=519;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=8234;
 RA Jacob-Dubuisson F., Kehoe B., Willery E., Reveneau N., Loch C.,
 RA Relman D.A.;
 RT "Molecular characterization of the Bordetella bronchiseptica
 RT filamentous hemagglutinin and its secretion machinery";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF111798; AAF21948.1; -;

FT NON_TER 18 18
 SQ SEQUENCE 18 AA: 2135 MW: F129A793B031E143 CRC64;

Query Match 23.5%; Score 24; DB 2; Length 18;
 Best Local Similarity 20.0%; Pred. No. 2e+03;
 Matches 3; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 1 SNMYAMMIRFKMP 15
 DB 3 TNLVRLVFSHVGMP 17

RESULT 7
 ID 035556 PRELIMINARY; PRT; 21 AA.
 AC 035556;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)
 DE HYPOTHETICAL 2.5 KDA PROTEIN.
 OS Pythium oligandrum.
 OC Mitochondrion.
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Pythium.
 NX NCBI_Taxid=41045;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17-1;
 RA Martin F.N.;
 RA Curr. Genet. 27:0-0(0).
 DR EMBL: U28355: AAA70034.1; -
 KW Hypothetical protein; Mitochondrion.
 SQ SEQUENCE 21 AA: 2491 MW: 883B7AB7A5D4445E CRC64;

Query Match 23.5%; Score 24; DB 8; Length 21;
 Best Local Similarity 36.4%; Pred. No. 2.3e+03;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 7 MIARFKMEPEV 17
 DB 1 MIVFKIYGD 11

RESULT 8
 ID 09J3U5 PRELIMINARY; PRT; 14 AA.
 AC 09J3U5;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE B-RAF PROTEIN (FRAGMENT).
 GN B-RAF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Barnier J.V., Papin C., Eyche A., Lecoc O., Calothy G.;
 RT "The mouse B-raf gene encodes multiple protein isoforms with tissue-specific expression";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ216308; CAB81556.1; -
 FT NON_TER 14 14
 FT NON_TER 14 14
 SQ SEQUENCE 14 AA: 1748 MW: D1E0505C44927F02 CRC64;

Query Match 22.5%; Score 23; DB 11; Length 14;
 Best Local Similarity 80.0%; Pred. No. 2.2e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 FPEVK 18
 DB 4 FPEVE 8

RESULT 9
 ID 085673 PRELIMINARY; PRT; 17 AA.
 AC 085673;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE REOVIRUS SEROTYPE 3 LI (FRAGMENT).
 OS Reovirus sp.
 OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
 NX NCBI_Taxid=10891;
 RN [1]
 RP SEQUENCE OF 1-6 FROM N.A.
 RX MEDLINE=82217029; PubMed=7086967;
 RA Kozak M.;
 RT "Sequences of ribosome binding sites from the large size class of reovirus mRNA";
 RL J. Virol. 42:467-473(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83017876; PubMed=6927854;
 RA Antczak J.B., Chmelo R., Pickup D.J., Joklik W.K.;
 RT "Sequence at both terminal of the 10 genes of reovirus serotype 3 (strain Dearling).";
 RL Virology 121:307-319(1982).
 DR EMBL: J02313: AAA47269.1; -
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA: 1929 MW: E36B60E1362BA8F1 CRC64;

Query Match 22.5%; Score 23; DB 12; Length 17;
 Best Local Similarity 35.7%; Pred. No. 2.7e+03;
 Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 3 MYAMMIRFKMEPE 16
 DB 1 MSSMILTFGPPIE 14

RESULT 10
 ID 000340 PRELIMINARY; PRT; 17 AA.
 AC 000340;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE HYPOTHETICAL PROTEIN IN O2L 3 REGION (FRAGMENT).
 OS Vaccinia virus (strain L-IVP).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 NX NCBI_Taxid=31531;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L-IVP;
 RA Ryzantkina O.I., Shchelkunov S.N., Muravlev A.I., Netesova N.A.,
 RA Mikryukov N.N., Gutovov V.V., Nikulin A.E., Kulichkov V.A.,
 RA Malugin E.G.;
 RT "Localization and determination of the nucleotide sequence of vaccinia virus genes coding for the proteins 36k and 12k";
 RL Dokl. Biochem. 247:780-787(1990).
 DR EMBL: X61166: CAB57393.1; -
 FT NON_TER 17 17
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA: 2017 MW: 848F2A871E90FBC CRC64;

Query Match 22.5%; Score 23; DB 12; Length 17;
 Best Local Similarity 50.0%; Pred. No. 2.7e+03;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 13 MPEVKER 20
 Db 3 MYEFARK 10

RESULT 11
 O9R4N5 PRELIMINARY; PRT; 18 AA.
 AC O9R4N5; 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE EXTRACTABLE ANTIGEN 1 (FRAGMENT).
 OS Bacillus anthracis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/staphylococcus group; Bacillus.
 OX NCBI_TaxID=1392;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95247684; PubMed=7730281;
 RA Farcaus J.W., Ribot W.J., Downs M.B., Ezzell J.W.;
 RT "Purification and characterization of the major surface array protein
 from the avirulent Bacillus anthracis Delta Sterne-1.";
 RL J. Bacteriol. 177:2481-2489(1995).
 SQ SEQUENCE 18 AA; 1926 MW; 1DBEBF0A4925EFB6 CRC64;

Query Match 22.5%; Score 23; DB 2; Length 18;
 Best Local Similarity 66.7%; Pred. No. 2.9e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 12 KMPEV 17
 Db 3 KTFPDV 8

RESULT 12
 O9R4W0 PRELIMINARY; PRT; 20 AA.
 AC O9R4W0; 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE 64 KRA HEAT SHOCK AND ALKALINE PH-INDUCED PROTEIN (FRAGMENT).
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95012609; PubMed=7927682;
 RA Wu Y.L., Lee L.H., Rollins D.M., Chung W.M.;
 RT "Heat shock and alkaline pH-induced proteins of Campylobacter jejuni:
 RT characterization and immunological properties.";
 RL Infect. Immun. 62:4256-4260(1994).
 SQ SEQUENCE 20 AA; 2339 MW; 0DDE08352969DEA7 CRC64;

Query Match 22.5%; Score 23; DB 2; Length 20;
 Best Local Similarity 45.5%; Pred. No. 3.2e+03;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 9 ARKMEPVKE 19
 Db 10 ARNKLYEGVKK 20

RESULT 13
 O05279 PRELIMINARY; PRT; 22 AA.
 AC O05279; 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
 DE HYPOTHETICAL 3.0 KDA PROTEIN.
 OS Chloranite-Aster yellow phytoplasma.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
 OC Acheilplasmataceae; Phytoplasma.
 OX NCBI_TaxID=54389;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Carder J.H., Barbara D.J., Davies D.L., Clarke M.F.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U77617; AAB51346.1; -
 KW Hypothetical protein
 SQ SEQUENCE 22 AA; 2995 MW; B14DBEC1453BB0F3 CRC64;

Query Match 22.5%; Score 23; DB 2; Length 22;
 Best Local Similarity 33.3%; Pred. No. 3.5e+03;
 Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 6 MFIARFMEPV 17
 Db 1 MFIYFLTFPOI 12

RESULT 14
 O9S551 PRELIMINARY; PRT; 22 AA.
 AC O9S551; 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE BETA-GALACTOSIDASE (FRAGMENT).
 GN LACZ.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Actinobacillus.
 OX NCBI_TaxID=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=4074;
 RC MEDLINE=99392473; PubMed=10463177;
 RA Schaller A., Kuhn R., Kuhnert P., Nicolet J., Anderson T.J.,
 RA MacInnes J.I., Segers R.P.A.M., Frey J.;
 RT "Characterization of aprIVA, a new RTX determinant of Actinobacillus
 RT pleuropneumoniae.";
 RL Microbiology 145:2105-2116(1999).
 DR EMBL: AF021919; AAD01699.1; -
 FT NON_TER 1
 SQ SEQUENCE 22 AA; 2849 MW; 762A86A651B6D3BB CRC64;

Query Match 22.5%; Score 23; DB 2; Length 22;
 Best Local Similarity 40.0%; Pred. No. 3.5e+03;
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 YAMIAEFKM 13
 Db 1 YRLMESEFKI 10

RESULT 15
 O62966 PRELIMINARY; PRT; 12 AA.
 ID O62966; 01-NOV-1996 (TREMBLrel. 01, Created)
 AC O62966; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 08, Last annotation update)
 DE INTERSTITIAL COLLAGENASE (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE-96408720; PubMed-8813727;
RA Rajakumar R.A., Quinn C.O.;
RT "Parathyroid hormone induction of rat interstitial collagenase mRNA in
  osteosarcoma cells is mediated through an AP-1-binding site.";
RL Mol. Endocrinol. 10:867-878(1996).
DR EMBL, U53605; AAB47407.1; -.
FT NON_TER
SQ SEQUENCE 12 AA; 1432 MW; 148A4DFE8ADD720 CRC64;

Query Match
Best Local Similarity 21.6%; Score 22; DB 11; Length 12;
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 MYAMMARFKMF 14
  |:::| | |
  1 MHSATLATFFFF 12

RESULT 16
ID Q9UMW1 PRELIMINARY; PRT; 15 AA.
AC Q9UMW1;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)
DE BENZOYL-COA LIGASE (FRAGMENT).
OS Methanospirillum hungatei.
OC Archaea; Euryarchaeota; Methanomicrobiales; Methanomicrobiaceae;
  Methanospirillum.
OC NCBI_TaxID-2203;
RN [1]
RP SEQUENCE.
RA Auburger G., Winter J.;
RL Appl. Microbiol. Biotechnol. 37:789-795(1992).
SQ SEQUENCE 15 AA; 1880 MW; D2972EF3B690AC5C CRC64;

Query Match
Best Local Similarity 21.6%; Score 22; DB 1; Length 15;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 KMFPE 16
  |:::| |
  2 KLYPE 6

RESULT 17
ID Q35921 PRELIMINARY; PRT; 15 AA.
AC Q35921;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-NOV-1998 (TRENBLREL. 08, Last annotation update)
DE MITOCHONDRIAL ATPASE 6 (FRAGMENT).
GN TRGELY.
OS Salmo salar (Atlantic salmon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
  Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OC NCBI_TaxID-8030;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE-95093530; PubMed-8000479;
RA Hardman G., Byrnes L., Peden J., Wolff J., Gannon F.;
RT "Cloning and sequencing of the Atlantic salmon (Salmo salar)
  cytochrome c oxidase subunit III gene (coxiii) and analysis of coxiii
  expression during parr-smolt transformation.";

```

```

RL Mol. Mar. Biol. Biotechnol. 3:210-216(1994).
DR EMBL, L04502; AAA62410.1; -.
KN Mitochondrion.
FT NON_TER
SQ SEQUENCE 15 AA; 1784 MW; 6BFA79B9E031C4BD CRC64;

Query Match
Best Local Similarity 21.6%; Score 22; DB 8; Length 15;
Matches 4; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 MYAMMARFKMF 15
  |:::| | |
  1 MIFLVLRMSDEP 13

RESULT 18
ID Q9UMJ9 PRELIMINARY; PRT; 16 AA.
AC Q9UMJ9;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
DE GALACTOSE-1-PHOSPHATE URIDYL TRANSFERASE (FRAGMENT).
GN GALT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID-9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Reichardt J.K.V.;
RT "The molecular genetic basis of galactosemia.";
RL Int. Pediatr. 8:110-113(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-95193790; PubMed-7887416;
RA Elias L.J., Langley S., Steele E., Evinger J., Fridovich-Kell J.L.,
  Brown A., Singh R., Fernhoff P., Hjelm L.N., Dembure P.P.;
RT "Galactosemia: a strategy to identify new biochemical phenotypes and
  molecular genotypes.";
RL Am. J. Hum. Genet. 56:630-639(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA Elias L.J., II, Langley S., Paulk E.M., Hjelm L.N., Dembure P.P.;
RT "A molecular approach to galactosemia.";
RL Eur. J. Pediatr. Suppl. 2:21-27(1995).
DR EMBL, L46718; AAB59602.1; -.
KN Transferase.
FT NON_TER
SQ SEQUENCE 16 AA; 1832 MW; 312B633F6A3D8A52 CRC64;

Query Match
Best Local Similarity 21.6%; Score 22; DB 4; Length 16;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 10 RPKFPEV 17
  |:::| | |
  4 RLRALPEV 11

RESULT 19
ID Q9RLP2 PRELIMINARY; PRT; 20 AA.
AC Q9RLP2;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
DE PUTATIVE CYSTEINYL T-RNA SYNTHETASE (FRAGMENT).
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OC NCBI_TaxID-485;

```

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WR302;
 RA Stein D.C., Gunn J.S.;
 RT "Use of a non-selective transformation technique to construct a
 RT multiple restriction-modification deficient mutant of *Neisseria*
 RT gonorrhoeae";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U42459; AAD56389.1; -.
 FT NON_TER 20
 SQ SEQUENCE 20 AA; 2245 MW; 4C893C704879FFD3 CRC64;

Query Match 21.6%; Score 22; DB 2; Length 20;
 Best Local Similarity 38.5%; Pred. No. 4.7e+03;
 Matches 5; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 7 MIAFKMPEVKE 19
 Db 1 MTAIYNTLPKKE 13

RESULT 20
 ID 09UK71 PRELIMINARY; PRT; 20 AA.
 AC 09UK71:
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
 DE LITHIUM-SENSITIVE MYO-INOSITOL MONOPHOSPHATASE A1 (EC 3.1.3.25)
 DE (FRAGMENT).
 GN IMPA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Parthasarathy L., Parthasarathy R.;
 RT "Molecular cloning, genomic organization and promoter analysis of the
 RT human brain lithium-sensitive myo-inositol monophosphatase A1
 RT isoenzyme";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF178754; AAD5297.1; -.
 DR HSSP; P29218; 11MB.
 KW Hydrolyase.
 FT NON_TER 20
 SQ SEQUENCE 20 AA; 2257 MW; 067C2976A73D64A4 CRC64;

Query Match 21.6%; Score 22; DB 4; Length 20;
 Best Local Similarity 57.1%; Pred. No. 4.7e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 YAMMIAR 10
 Db 11 YAVTIAR 17
 RESULT 21
 ID 029341 PRELIMINARY; PRT; 20 AA.
 AC 029341:
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, last annotation update)
 DE NGF-INDUCIBLE PROTEIN TIS21 (FRAGMENT).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE-SMALL INTESTINE;
 RA Minireoe A.K., Fredholm M., Davies W.;
 RL Mamm. Genome 7:509-517(1996).
 DR EMBL; F14751; CAA23227.1; -.
 FT NON_TER 1
 SQ SEQUENCE 20 AA; 2058 MW; DEC82EC21D641A0 CRC64;

Query Match 21.6%; Score 22; DB 6; Length 20;
 Best Local Similarity 50.0%; Pred. No. 4.7e+03;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 10 RFKMFPEV 17
 Db 3 RDMPEI 10

RESULT 22
 ID 09TRCO PRELIMINARY; PRT; 20 AA.
 AC 09TRCO:
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
 DE ENTEROTOXIN-BINDING GLYCOPROTEIN P16K (FRAGMENT).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=94259890; PubMed=8201051;
 RA Shida K., Takamizawa K., Nagaoka M., Kushi A., Osawa T., Tsuji T.;
 RT "Enterotoxin-binding glycoproteins in a protease-peptone fraction of
 RT heated bovine milk";
 RL J. Dairy Sci. 77:930-939(1994).
 DR HSSP; P00711; 1HFZ.
 SQ SEQUENCE 20 AA; 2321 MW; C738FD14F55C74C3 CRC64;

Query Match 21.6%; Score 22; DB 6; Length 20;
 Best Local Similarity 25.0%; Pred. No. 4.7e+03;
 Matches 3; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 8 IARFKMPEVKE 19
 Db 3 LTKXEVFRELKD 14

RESULT 23
 ID 09XSLS PRELIMINARY; PRT; 21 AA.
 AC 09XSLS:
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, last annotation update)
 DE ALPHA S2-CASEIN (FRAGMENT).
 GN CSNIS2.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Veltri C., Pilla F., Lagonigro R.;
 RT "A new allele of goat alpha s2-casein";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ242728; CAB44433.2; -.
 FT NON_TER 1
 SQ SEQUENCE 21

SO SEQUENCE 21 AA: 2591 MW: A608995532CA3098 CRC64;

Query Match 21.6%; Score 22; DB 6; Length 21;

Best Local Similarity 33.3%; Pred. No. 4.9e+03; Mismatches 3; Indels 0; Gaps 0;

OY 8 IAREKMEPEVK 19
10 ISRRKIMPEIPE 21

RESULT 24

O92E05 PRELIMINARY; PRT; 24 AA.

AC O92E05; PRELIMINARY; PRT; 24 AA.

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE PORA PROTEIN (FRAGMENT).

GN PORA.

OS Neisseria meningitidis.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

OY NCBI_TaxID=487;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-19/92;

RA Wedge E., Caugant D.A., Musacchio A., Saunders N.B., Zollinger W.D.;

RT *Specificities of reference monoclonal antibodies and human

RT postvaccination sera with meningococcal P1.19,15 outer membrane

RT proteins.

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ012727; CA10150.1; -

FT NON_TER 1 24

FT NON_TER 1 24

SO SEQUENCE 24 AA: 2743 MW: 260BA002923D9F6 CRC64;

Query Match 21.6%; Score 22; DB 2; Length 24;

Best Local Similarity 55.6%; Pred. No. 5.6e+03; Mismatches 3; Indels 0; Gaps 0;

OY 10 RFKMEPEVK 18
10 RSKSQPOVK 18

RESULT 25

O9XBK2 PRELIMINARY; PRT; 24 AA.

AC O9XBK2; PRELIMINARY; PRT; 24 AA.

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)

DE TRANSLATION ELONGATION FACTOR TS (EF-TS) (FRAGMENT).

GN TSF.

OS Bacillus cereus.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OY NCBI_TaxID=1396;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 10987;

RX MEDLINE-99231848; PubMed-10217496;

RT *Genome organization is not conserved between Bacillus cereus and

RT Bacillus subtilis.

RL Microbiology 145:621-631(1999).

DR EMBL: AJ010135; CAB40592.1; -

FT NON_TER 24

SO SEQUENCE 24 AA: 2671 MW: F1AB99BFEB5DDABD CRC64;

Query Match 21.6%; Score 22; DB 2; Length 24;

Best Local Similarity 44.4%; Pred. No. 5.6e+03; Mismatches 2; Indels 0; Gaps 0;

OY 12 KMEPEYK 20
6 QMVKELREK 14

Search completed: February 5, 2001, 10:55:22
Job time: 911 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:47:29 ; Search time 68.99 Seconds

(without alignments)
9.913 Million cell updates/sec

Title: US-08-981-824-2

Perfect score: 102

Sequence: 1 SNMTAMTARFKEPEVKEK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36.*

1: /SIDSI/gcgdata/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	20	18	W18843
2	102	100.0	20	18	W01794
3	98	96.1	20	16	R72277
4	98	96.1	20	21	Y59541
5	98	96.1	20	21	Y59577
6	78	76.5	20	18	W10300
7	74	72.5	14	16	R76653
8	74	72.5	14	18	W18861
9	74	72.5	23	13	R29627
10	66	64.7	13	18	W35531
11	53	52.0	24	13	R29626
12	37	36.3	9	20	Y30061

13	32	31.4	15	21	Y09016
14	32	31.4	19	13	R27147
15	31	30.4	11	18	W10301
16	31	30.4	20	16	R72276
17	31	30.4	20	16	R72278
18	31	30.4	20	21	Y59576
19	31	30.4	20	21	Y59578
20	31	30.4	20	21	Y57065
21	30	29.4	14	14	R43871
22	30	29.4	15	17	R66007
23	30	29.4	22	11	R06363
24	30	29.4	25	19	W63838
25	30	29.4	25	19	W63841
26	29	28.4	14	18	W40976
27	29	28.4	16	9	P82903
28	29	28.4	19	17	R96938
29	29	28.4	19	17	R96940
30	29	28.4	19	17	R96941
31	28.5	27.9	15	20	Y30477
32	28	27.5	10	7	P60768
33	28	27.5	19	11	Y07348
34	28	27.5	19	17	R96939
35	27	26.5	9	20	Y46838
36	27	26.5	11	16	R82084
37	27	26.5	11	17	R99509
38	27	26.5	11	18	W12233
39	27	26.5	11	19	W60689
40	27	26.5	13	10	P90569
41	27	26.5	13	20	W93458
42	27	26.5	14	18	W40975
43	27	26.5	14	18	W40978
44	27	26.5	14	18	W40979
45	27	26.5	14	18	W40980

ALIGNMENTS

RESULT 1

W18843 standard; peptide: 20 AA.

AC W18843;

DT 05-JAN-1998 (first entry)

DE 65 KD Glutamic acid decarboxylase peptide fragment II.

XX

KW GAD, 65 KD; human; glutamic acid decarboxylase; autoreactive; diagnosis;

KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;

KW predisposition; autoimmune; tumour; rheumatoid arthritis;

KW multiple sclerosis.

OS Synthetic.

XX

PN DE19526561-A1.

PD 23-JAN-1997.

PF 20-JUL-1995; 95DE-1026561.

PR 20-JUL-1995; 95DE-1026561.

XX

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;

PI Pezzilli P, Stahl P;

XX WPI; 1997-088254/09.

XX

PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -

XX Involving intradermal admin. of auto-reactive substances

HLA class II bindi

Proinsulin express

Mammalian GAD 65 p

Glutamic acid deca

Glutamic acid deca

GAD65 fragment, pe

GAD65 fragment, pe

Glutamate decarbox

OmpA2-R-7 signal p

Nuclear localisati

C-terminal of fish

Plasmod pmp100 M.

Plasmod pmp100 M.

Cryptic peptide of

Activated metallopro

Loop swap I mutain

Loop swap IIL mutain

LS-Pro mutain of a

Mutant sequence of

Sequence of peptide

Matrix metallopro

Loop swap II mutain

Immunogenic peptide

Influenza hemagglut

T-cell epitope use

Influenza PR8 A VI

Influenza PR8A vir

Peptide to increas

Antigenic peptide

Cryptic peptide of

Cryptic peptide of

Cryptic peptide of

PS Claim 11; Page 9; 12pp; German.
 XX
 CC W18842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predisposition to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 XX
 SO Sequence 20 AA;
 Query Match 100.0%; Score 102; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNMYAMTARFKMPEVKEK 20
 DB 1 smymammarfkmpkevkek 20
 RESULT 2
 ID W01794 standard; peptide; 20 AA.
 XX W01794;
 XX 15-OCT-1997 (first entry)
 DE Human 65 kD glutamine decarboxylase peptide.
 KW Human; glutamine decarboxylase; GAD; diagnosis; predisposition;
 KW tumour; immunological; disease; autoimmune; diabetes; reagent;
 KW determination; T cell; subpopulation; medicament; treatment;
 KW prevention; production; antigen; immunogen; tolerogen; isolation;
 KW reinjection; inactivation.
 XX Homo sapiens.
 OS
 PN DE19525784-A1.
 XX 16-JAN-1997.
 PD 14-JUL-1995; 95DE-1025784.
 XX 14-JUL-1995; 95DE-1025784.
 PF 14-JUL-1995; 95DE-1025784.
 XX 14-JUL-1995; 95DE-1025784.
 PR (BOE) BOEHRINGER MANNHEIM GMBH.
 PA
 XX (BOE) BOEHRINGER MANNHEIM GMBH.
 PI Albert W, Boltard C, Endl J, Jung G, Schendel D;
 PI Stahl P, Van Endert P;
 XX
 DR WPI: 1997-078452/08.
 XX
 PT Glutamine decarboxylase peptide(s) - for diagnosis and therapy of
 PT diabetes, etc.
 XX
 PS Claim 1; Page 12; 15pp; German.
 XX
 CC The present peptide is a fragment of the human 65 kD glutamine
 CC decarboxylase (GAD), which can be used to diagnose, or diagnose a
 CC predisposition to, a tumour or immunological disease, preferably an
 CC autoimmune disease, especially diabetes. It can also be used as a
 CC reagent to determine specific T cell subpopulations, in medicaments
 CC to treat or prevent immunological diseases, preferably autoimmune
 CC diseases, especially diabetes; to produce antigens, especially
 CC immunogens or tolerogens and to isolate specific T cell
 CC subpopulations, which can be used to produce antigens or for
 CC reinjection, optionally after inactivation.

XX
 SO Sequence 20 AA;
 Query Match 100.0%; Score 102; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNMYAMTARFKMPEVKEK 20
 DB 1 smymammarfkmpkevkek 20
 RESULT 3
 ID R72277 standard; Peptide; 20 AA.
 XX R72277;
 AC R72277;
 XX 13-NOV-1995 (first entry)
 DE Glutamic acid decarboxylase (GAD65) fragment.
 XX
 KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
 KW insulin-dependant diabetes mellitus; stiff man disease.
 XX
 OS Homo sapiens.
 PN W09507992-A.
 XX 23-MAR-1995.
 PD 24-AUG-1994; 94WO-US09478.
 XX 17-SEP-1993; 93US-0123859.
 PR (REGC) UNIV CALIFORNIA.
 PA
 XX Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
 PI WPI: 1995-131360/17.
 DR
 XX
 PT New polypeptide fragments of glutamic acid decarboxylase - for
 PT diagnosis and treatment of autoimmune disease, esp. insulin
 PT dependent diabetes, also related nucleic acid, vectors,
 PT antibodies, hydridoma(s) etc.
 XX
 PS Claim 1; Page 76; 100pp; English.
 XX
 CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
 CC acid decarboxylase (GAD65) respectively, from which the GAD65
 CC fragments described in R72261-R72298 were derived. These fragments
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose
 CC and treat GAD-related autoimmune disorders, such as insulin
 CC dependant diabetes mellitus or stiff man disease.
 XX
 SO Sequence 20 AA;
 Query Match 96.1%; Score 98; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 NMVAMTARFKMPEVKEK 20
 DB 1 nmymammarfkmpkevkek 19
 RESULT 4
 ID Y59541 standard; peptide; 20 AA.
 XX Y59541;
 AC Y59541;

```

XX 03-APR-2000 (first entry)
DT GAD65 fragment, peptide #17.
XX
DE GAD65 fragment, peptide #17.
XX
KM GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
KW insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
therapy.
XX
XX Homo sapiens.
OS
XX US5998366-A.
PN
XX 07-DEC-1999.
PD
XX 09-APR-1997; 97US-0827618.
PF
XX 07-JUN-1995; 95US-0485725.
PR 21-SEP-1990; 90US-0586536.
PR 18-JUN-1991; 91US-0716909.
XX
XX (REGC ) UNIV CALIFORNIA.
PA
PI Tobin AJ, Kaufman DL, Erlander MG;
XX
DR WPI; 2000-095930/08.
XX
PT Ameliorating glutamic acid decarboxylase associated autoimmune
PT disorders such as insulin dependent diabetes mellitus and Stiff man
PS disease -
XX
PS Claim 1; Column 33; 61pp; English.
XX
CC This sequence represents a fragment of the glutamic acid decarboxylase 65
CC (GAD65) protein. The invention relates to a method of ameliorating GAD
CC associated autoimmune disorders by administering a GAD65 peptide to the
CC patient. The method can be used for ameliorating GAD associated
CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
CC and Stiff man disease. GAD65 can also be useful for screening drugs that
CC alter GAD function, for generating monoclonal antibodies and in
CC immunosays. GAD65 is an effective diagnostic tool for predicting IDDM
CC and the diagnosis is quite easy. It is also possible to obtain much
CC larger quantities of polypeptide via recombinant techniques than are
CC available from natural sources.
XX
SQ Sequence 20 AA;

```

```

Query Match          96.1%; Score 98; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 NMYAMMIARFKMPEVKEK 20
   |||||
Db 1 nmyammiarfkmpfvkex 19

```

```

RESULT 5
ID Y59577 standard; peptide; 20 AA.
XX
AC Y59577;
XX
DT 03-APR-2000 (first entry)
XX
DE GAD65 fragment, peptide #17.
XX
KM GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
KW insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
therapy.
XX
XX Homo sapiens.
OS
XX

```

```

PN US5998366-A.
XX
XX 07-DEC-1999.
PD
XX 09-APR-1997; 97US-0827618.
PF
XX 07-JUN-1995; 95US-0485725.
PR 21-SEP-1990; 90US-0586536.
PR 18-JUN-1991; 91US-0716909.
XX
XX (REGC ) UNIV CALIFORNIA.
PA
PI Tobin AJ, Kaufman DL, Erlander MG;
XX
DR WPI; 2000-095930/08.
XX
PT Ameliorating glutamic acid decarboxylase associated autoimmune
PT disorders such as insulin dependent diabetes mellitus and Stiff man
PS disease -
XX
PS Example 11; Column 42; 61pp; English.
XX
XX This sequence represents a fragment of the glutamic acid decarboxylase 65
XX (GAD65) protein. The invention relates to a method of ameliorating GAD
XX associated autoimmune disorders by administering a GAD65 peptide to the
XX patient. The method can be used for ameliorating GAD associated
XX autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
XX and Stiff man disease. GAD65 can also be useful for screening drugs that
XX alter GAD function, for generating monoclonal antibodies and in
XX immunosays. GAD65 is an effective diagnostic tool for predicting IDDM
XX and the diagnosis is quite easy. It is also possible to obtain much
XX larger quantities of polypeptide via recombinant techniques than are
XX available from natural sources.
XX
SQ Sequence 20 AA;

```

```

Query Match          96.1%; Score 98; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 NMYAMMIARFKMPEVKEK 20
   |||||
Db 1 nmyammiarfkmpfvkex 19

```

```

RESULT 6
ID W10300 standard; peptide; 20 AA.
XX
AC W10300;
XX
DT 11-SEP-1997 (first entry)
XX
DE Mammalian GAD 65 peptide.
XX
KW Soluble; fusion; major histocompatibility complex; MHC;
KW heterodimer; complex; GAD 65; antigen; binding groove; tolerance;
KW autoantigen; disease; insulin dependent; diabetes mellitus; IDDM;
KW antagonist; T cell; energy; presenting cell.
XX
OS Mammalian spp.
XX
XX WO9640944-A2.
PN
XX 19-DEC-1996.
PD
XX 07-JUN-1996; 96WO-US10102.
PF
XX 27-OCT-1995; 95US-0005964.
PR 07-JUN-1995; 95US-0480002.
PR 07-JUN-1995; 95US-0482133.
PR 07-JUN-1995; 95US-0483241.

```

XX (ANER-) ANERGEN INC.
 PA (ZYMO) ZYMOGENETICS INC.
 XX Deshpande S, Gross JA, Kindsvogel W, Reich EP, Sheppard PO;
 PI WPI; 1997-05237/05.
 XX
 XX Novel fused major histocompatibility complex:antigenic peptide
 PT complex - useful to induce tolerance to an autoantigen-related
 PR disease e.g. insulin-dependent diabetes mellitus
 XX
 PS Claim 10; Page 110; 142pp; English.
 XX
 CC A novel soluble fused major histocompatibility complex (MHC)
 CC heterodimer:peptide complex, comprises DNA encoding 1st and 2nd
 CC MHC domains, linked by DNA encoding a 5-25 residue linker, and a
 CC DNA encoding an antigenic peptide able to associate with a peptide
 CC binding groove of the MHC molecule, e.g. the present peptide,
 CC linked in frame to the DNA encoding the 2nd domain by a DNA
 CC encoding a 5-25 residue linker. The complex can be used to induce
 CC immunological tolerance in adults susceptible to, or suffering from
 CC an autoimmune related disease, e.g. insulin dependent diabetes
 CC mellitus (IDDM), by antagonising the binding of particular T cells
 CC and antigen presenting cells, to induce energy (immunological
 CC non-responsiveness) in the targeted T cell. As the heterodimers and
 CC corresponding antigen are permanently linked into a single chain,
 CC obviating the requirement for complex heterodimer truncation or
 CC formation, the complex eliminates inefficient and non-specific
 CC peptide loading.
 XX
 SO Sequence 20 AA;

Query Match 76.5%; Score 78; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.1e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNMYAMTARFKMP 15
 DB 6 snmyamtarfkmp 20

RESULT 7
 R76653
 ID R76653 standard; peptide; 14 AA.
 XX
 AC R76653;
 XX
 DT 05-MAR-1996 (first entry)
 XX
 DE Peptide derived from human glutamic acid decarboxylase 12.
 XX
 KW diabetes; T-cell subpopulation; detection; antigen production;
 KW diagnosis; autoimmune disease.
 XX
 OS Homo sapiens.
 XX
 PN DE4418091-A1.
 PD 27-JUL-1995;
 XX
 PE 24-MAY-1994; 94DE-4418091.
 XX
 PR 04-FEB-1994; 94DE-4403522.
 PR 20-JAN-1994; 94DE-4401629.
 XX
 PA (ENDL/) ENDL J.
 PA (BOE) BOEHRINGER MANNHEIM GMBH.
 XX
 PI Albert W, Dormair K, Endl J, Jung G, Meinel E;
 PI Stahl P, Schendel D;
 XX

DR WPI; 1995-264505/35.
 XX
 XX Antigen-specific activated T-lymphocytes and their detection - by
 PT interaction with inventive peptide(s) of peptide-MHC complexes;
 PR useful in diagnosis of e.g. diabetes and autoimmune diseases
 XX
 PS Claim 1; Fig 2; 21pp; German.
 XX
 CC R76642-62 are derived from human glutamic acid decarboxylase and
 CC specifically react with T-cell sub-populations isolated from recently
 CC diagnosed Type-1 diabetics. Pharmaceutical compans. conty. these
 CC peptides and those shown in R77571-72, are useful for the diagnosis of
 CC a disease or predispositions of immune system diseases, tumours, and
 CC autoimmune diseases, including diabetes. The peptides are able to detect
 CC specific T-cell subpopulations that are then used for antigen prodn.,
 CC e.g. by reinjection.
 XX
 SO Sequence 14 AA;

Query Match 72.5%; Score 74; DB 16; Length 14;
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NMVAMTARFKMP 15
 DB 1 nmvamtarfkmp 14

RESULT 8
 W18861
 ID W18861 standard; peptide; 14 AA.
 XX
 AC W18861;
 XX
 DT 05-JAN-1998 (first entry)
 XX
 DE 65 kD glutamic acid decarboxylase peptide fragment 12.
 XX
 KW GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
 KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
 KW predisposition; autoimmune; tumour; rheumatoid arthritis;
 KW multiple sclerosis.
 XX
 OS Synthetic.
 XX
 PN DE19526561-A1.
 PD 23-JAN-1997.
 XX
 PE 20-JUL-1995; 95DE-1026561.
 XX
 PR 20-JUL-1995; 95DE-1026561.
 XX
 PA (BOE) BOEHRINGER MANNHEIM GMBH.
 XX
 PI Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;
 PI Pozzilli P, Stahl P;
 PI WPI; 1997-088254/09.
 DR
 XX
 PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
 PT involving intradermal admin. of auto-reactive substances
 XX
 PS Claim 11; Fig 2; 12pp; German.
 XX
 CC W18842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predisposition to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive

CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 XX

SO Sequence 14 AA;

Query Match 72.5%; Score 74; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 NMYAMMARRKMF 15
 |||||
 Db 1 nmyammarrkmp 14

RESULT 9
 R29627
 ID R29627 standard: peptide; 23 AA.

AC R29627;

DT 05-MAY-1993 (first entry)

DE GAD peptide.

XX Glutamic acid decarboxylase; auto immune disease; insulin dependent;
 KW diabetes mellitus; drug screening; antibody; stiff man syndrome.
 XX

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 15 /label= Thr, Glu

PN EP519469-A.

PD 23-DEC-1992.

PF 17-JUN-1992; 92EP-0110308.

PR 18-JUN-1991; 91US-0716909.

XX (REGC) UNIV CALIFORNIA.

XX Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;

DR WPI; 1992-425701/52.

PT Glutamic acid decarboxylase isolated polypeptide(s) - useful for
 PT diagnosis and treatment of auto-immune diseases, e.g. insulin
 PT dependent diabetes mellitus, drug screening and antibody prodn.

XX Claim 6; Page 20; 35pp; English.

CC The peptide corresponds to a glutamic acid decarboxylase (GAD 65)
 CC sequence and may be used for the detection of autoantibodies to GAD
 CC 65. They can also be used for screening drugs such as those that
 CC alter GAD function and for generation of antibodies which can be
 CC used to detect GAD 65. Such methods can be used in the diagnosis
 CC and therapy of autoimmune diseases such as insulin dependent
 CC diabetes mellitus and "stiff man" syndrome. The polypeptide can be
 CC used to bind or block the continuance of an autoimmune response to
 CC GAD 65. It may also be coupled to therapeutic agents and used
 CC immunotherapeutically. See also R29625-9.

XX Sequence 23 AA;

Query Match 72.5%; Score 74; DB 13; Length 23;
 Best Local Similarity 93.8%; Pred. No. 5.6e-07;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 AMMIARRKMFPEVK 20
 |||||
 Db 1 ammiarrkmpfvkxk 16

Query Match 64.7%; Score 66; DB 18; Length 13;
 Best Local Similarity 100.0%; Pred. No. 8.5e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 IARRKMFPEVK 20
 |||||
 Db 1 iarrkmpfvkxk 13

RESULT 10
 W35531
 ID W35531 standard: peptide; 13 AA.

AC W35531;

DT 22-APR-1998 (first entry)

DE Glutamic acid decarboxylase peptide GAD65 253-265.

XX T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
 KW scaffold; inhibition; metastasis; wound healing; solid phase.
 XX

OS Unidentified.

XX WO9738011-A1.

PD 16-OCT-1997.

PF 03-APR-1997; 97WO-DE00146.

PR 03-APR-1996; 96DK-0000398.

XX (PEPR-) PEPRSEARCH AS.

XX Heegaard PMH, Jakobsen PH;

DR WPI; 1997-512645/47.

XX Non-dendritic peptide carrier linked to a solid phase - useful as a
 XX diagnostic agent and as a scaffold for production of chemical
 XX derivatives

XX Example 27; Page 150; 262pp; English.

CC A non-dendritic peptide carrier (A) has been developed which is coupled
 CC through a linker to a solid phase, forming a complex of (A)-solid phase.
 CC where (A) comprises 10-50 amino acids capable of forming a secondary
 CC structure in a benign buffer after liberation from the solid phase, and
 CC further the (A)-solid phase complex comprises an immunogenic substance
 CC and/or an immune mediator coupled on (A). The present sequence
 CC represents a peptide used in an example from the present invention. An
 CC (A)-solid phase complex can be used as a scaffold for the production of
 CC chemical derivatives, characterised by covalently attaching molecules at
 CC attachment points. Alternatively (A) is used as a scaffold-peptide for
 CC the incorporation into an immunostimulating Complex (Iscom) resulting an
 CC (A)-Iscom complex which is used for the chemical coupling of antigenic
 CC substances in an aqueous solution by conjugation. (A) derivatised with
 CC one or more peptides having fibronectin-, laminin- or vitronectin-like
 CC binding activities can be used for the promotion of cell-attachment to
 CC plastic surfaces, in particular to inhibit tumour growth and metastasis,
 CC and for promotion of wound healing. Also a derivatised (A) can be used
 CC for the selection of specifically-binding aptamers or as a diagnostic
 CC agent. Such diagnostic (A) molecules could be used to detect molecules
 CC derived from or indicative of pregnancy or of a disease, such as an
 CC infectious, autoimmune or cancerous disease.

XX Sequence 13 AA;

Query Match 64.7%; Score 66; DB 18; Length 13;
 Best Local Similarity 100.0%; Pred. No. 8.5e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

XX AC Y99016;
XX
XX 07-AUG-2000 (first entry)
XX
XX HLA class II binding antigen epitope peptide #205.
DE
XX
XX Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;
XX immune response; chronic viral disease; cancer; autoimmune disease;
XX rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;
XX allograft rejection; allergy; Lyme disease; hepatitis; prostate cancer;
XX glomerulonephritis; food hypersensitivity; malaria.
XX
XX Unidentified.
OS
XX WO9961916-A1.
XX
XX 02-DEC-1999.
XX
XX 28-MAY-1999; 99WO-US12066.
XX
XX 29-MAY-1998; 98US-0087192.
XX
XX (EPIM-) EPIMUNE INC.
XX
XX Sette A, Southwood S, Sidney J;
XX
XX WPI: 2000-097143/08.
XX
XX New compositions containing immunogenic peptide epitopes for various
XX HLA class II DR molecules useful for inducing helper T cell response
XX
XX Claim 1: Page 43; 60pp; English.
XX
XX The present invention relates to a new pharmaceutical composition
XX comprising a unit dose form of a peptide, or analogue, comprising an
XX epitope selected from those represented by peptides Y9812-Y9939 which
XX are derived from various antigens for various human leucocyte antigen
XX class DR molecules, representative of the world wide population. The
XX peptide/analogue binds to an HLA class II molecule at an IC-50 of less
XX than or equal to 1,000 nM. The pharmaceutical can be used to induce a
XX helper T cell response. The pharmaceutical focuses the immune response
XX towards selected determinants and could therefore be used in cases of
XX chronic viral diseases and cancer. Examples of diseases that can be
XX treated using the peptide containing pharmaceutical include autoimmune
XX diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia
XX gravis), allograft rejection, allergies, Lyme disease, hepatitis,
XX post-streptococcal endocarditis or glomerulonephritis and food
XX hypersensitivities. The peptide epitopes can be used to enhance immune
XX responses against other immunogens administered with the peptides.
XX Diseases which can be treated using immunogenic mixtures include prostate
XX cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical
XX carcinoma, lymphoma, and condyloma acuminatum. The peptides may also be
XX used to make monoclonal antibodies useful as potential diagnostic or
XX therapeutic agents. The peptides may also be useful as diagnostic
XX reagents, for example, to determine the susceptibility of an individual
XX to a treatment regimen. Also, the peptides may be used to predict which
XX individuals will be at substantial risk of developing chronic infection.
XX The selection of appropriate T and B cell epitopes should allow the
XX development of epitope based vaccines particularly towards conserved
XX epitopes of pathogens which are characterized by high sequence
XX variability such as HIV, HCV and Malaria.
XX
XX Sequence 15 AA:
SQ

```

```

Query Match 31.4%; Score 32; DB 21; Length 15;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 10 RFKMPPEVKE 19
   11:1111111111
Db 1 rfemfrelne 10

```

```

RESULT * 14
R27147
ID R27147 standard; Protein: 19 AA.
XX
XX R27147;
XX
XX 20-MAY-1998 (first entry)
XX
XX Proinsulin expression stabilising leader sequence.
XX
XX Enhanced expression; foreign gene stabilisation.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FH 1..5
XX FT region /note= "beta-Gal"
XX FT region 6..17
XX FT region /note= "selected leader sequence - (Leu)6"
XX FT region 18..19
XX FT /note= "proinsulin start"
XX
XX CA1306208-C.
XX
XX 11-AUG-1992.
XX
XX 09-JAN-1985; 85CA-0471799.
XX
XX 09-JAN-1985; 85CA-0471799.
XX
XX (CANA-) CANADA NAT RES COUNCIL.
XX (WUR/) WU R J.
XX
XX Narang SA, Sung WL, Wu RJ;
XX
XX WPI: 1992-308288/38.
XX
XX N-PSDB; Q28661.
XX
XX Enhanced expression and/or stabilisation of foreign gene prods.
XX using short selected leader signal sequences in host microorganisms
XX
XX Example; Page 6; 17pp; English.
XX
XX The sequence comprises a short leader signal sequence (Leu)6 selected
XX to enhance the expression of and/or stabilise the proinsulin when
XX expressed in a host cell. This method can be used for the prodn. of
XX other foreign gene products such as insulin, somatostatin or an
XX interferon. See also R27146.
XX
XX Sequence 19 AA:
SQ

```

```

Query Match 31.4%; Score 32; DB 13; Length 19;
Best Local Similarity 28.6%; Pred. No. 19;
Matches 4; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 SNTYAMIRAFKMF 14
   11:1111111111
Db 5 tnl1111111qfmrnf 18

```

```

RESULT 15
W10301
ID W10301 standard; peptide: 11 AA.
XX
XX W10301;
XX
XX 11-SEP-1997 (first entry)
XX
XX Mammalian GAD 65 peptide.
XX

```

KM	Soluble fusion; major histocompatibility complex; MHC;
KW	heterodimer; complex; GAD 65; antigen; binding groove; tolerance;
KV	autoantigen; disease; insulin dependent; diabetes mellitus; IDDM;
KX	antagonist; T cell; anergy; presenting cell.
XX	
OS	Mammalian spp.
PN	MO9640944-A2.
PD	19-DEC-1996.
PF	07-JUN-1996; 96WO-US10102.
PR	27-OCT-1995; 95US-0005964.
PR	07-JUN-1995; 95US-0480002.
PR	07-JUN-1995; 95US-0482133.
PR	07-JUN-1995; 95US-0483241.
XA	(ANER-) ANERGEN INC.
PA	(ZYMO) ZYMOGENETICS INC.
XX	
PI	Deshpande S, Gross JA, Kindsvogel W, Reich EP, Shepard PO;
DR	WPI: 1997-052337/05.
XX	
PT	Novel fused major histocompatibility complex:antigenic peptide
PT	complex - useful to induce tolerance to an autoantigen-related
PS	disease e.g. insulin-dependent diabetes mellitus
XX	
PS	Claim 10; Page 109; 142pp; English.
XX	
CC	A novel soluble fused major histocompatibility complex (MHC)
CC	heterodimer:peptide complex, comprises DNA encoding 1st and 2nd
CC	MHC domains, linked by DNA encoding a 5-25 residue linker, and a
CC	DNA encoding an antigenic peptide able to associate with a peptide
CC	binding groove of the MHC molecule, e.g. the present peptide,
CC	linked in frame to the DNA encoding the 2nd domain by a DNA
CC	encoding a 5-25 residue linker. The complex can be used to induce
CC	immunological tolerance in adults susceptible to, or suffering from
CC	an autoimmune related disease, e.g. insulin dependent diabetes
CC	mellitus (IDDM), by antagonising the binding of particular T cells
CC	and antigen presenting cells, to induce anergy (immunological
CC	non-responsiveness) in the targeted T cell. As the heterodimers and
CC	corresponding antigen are permanently linked into a single chain,
CC	obviating the requirement for complex heterodimer truncation or
CC	formation, the complex eliminates inefficient and non-specific
CC	peptide loading.
XX	
SC	Sequence 11 AA:
QY	Query Match 30.4%; Score 31; DB 18; Length 11;
ID	Best Local Similarity 100.0%; Pred. No. 16;
XX	Matches 6; Conservative 0; Indels 0; Gaps 0;
OY	1 SNMYAM 6
DB	
	6 snmyam 11
RESULT 16	
R72276	R72276 standard; Peptide; 20 AA.
AC R72276;	
DT 13-NOV-1995	(first entry)
DE Glutamic acid decarboxylase (GAD65) fragment.	
XX Glutamic acid decarboxylase; GAD65; autoimmune disorders;	
KW insulin-dependent diabetes mellitus; stiff man disease.	
XX	

OS	Homo sapiens.
XX	
PX	W09507992-A.
PD	23-MAR-1995.
XX	
PX	24-AUG-1994;
PF	94WO-USO9478.
XX	
PR	17-SEP-1993; .
XX	93US-0123859.
PA	(RESC) UNITV CALIFORNIA.
XX	
PI	Clare-Salpler MJ, Erlander MG,
XX	Kaufman DL, Tobin AJ;
DR	WPI: 1995-131360/17.
XX	
PT	New polypeptide fragments of glutamic acid decarboxylase - for
PT	diagnosis and treatment of auto-immune disease, esp. insulin
PT	dependent diabetes, also related nucleic acid, vectors,
PT	antibodies, hybridoma(s) etc.
XX	
PS	Example 11; Page 76; 100pp; English.
CC	
XX	Q086481 and Q086482 encode R71733 and R79105, rat and human glutamic
CC	acid decarboxylase (GAD65) respectively, from which the GAD65
CC	fragments described in R72261-R72298 were derived. These fragments
CC	can be used to detect autoantibodies against GAD, e.g. to diagnose
CC	and treat GAP-related autoimmune disorders, such as insulin
CC	dependent diabetes mellitus or stiff man disease.
XX	
SQ	Sequence 20 AA:
Query Match	30.4%; Score 31; DB 16; Length 20;
Best Local Similarity	100.0%; Pred. No. 30;
Matches 6;	Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY	1 SNNYAM 6 Db 15 snmyam 20
RESULT 17	
ID	R72278
AC	R72278 standard; Peptide; 20 AA.
XX	
AC	R72278;
XX	
DY	13-NOV-1995 (first entry)
XX	
DE	Glutamic acid decarboxylase (GAD65) fragment.
XX	
KM	Glutamic acid decarboxylase; GAD65; autoimmune disorders;
KW	Insulin-dependant diabetes mellitus; stiff man disease.
XX	
OS	Homo sapiens.
XX	
PX	W09507992-A.
PN	
PD	23-MAR-1995.
XX	
PX	24-AUG-1994;
PF	94WO-USO9478.
XX	
PR	17-SEP-1993; .
XX	93US-0123859.
PA	(RESC) UNITV CALIFORNIA.
XX	
PI	Clare-Salpler MJ, Erlander MG, Kaufman DL, Tobin AJ;
XX	
DR	WPI: 1995-131360/17.
XX	
PT	New polypeptide fragments of glutamic acid decarboxylase - for
PT	diagnosis and treatment of auto-immune disease, esp. insulin
PT	dependent diabetes, esp. Insulin

PT dependent diabetes, also related nucleic acid, vectors,
PT antibodies, hydridoma(s) etc.
XX
PS Example 11; Page 76; 100pp; English.
XX
CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
CC acid decarboxylase (GAD65) respectively, from which the GAD65
CC fragments described in R72261-R72298 were derived. These fragments
CC can be used to detect autoantibodies against GAD, e.g. to diagnose
CC and treat GAD-related autoimmune disorders, such as insulin
CC dependant diabetes mellitus or stiff man disease.
XX
SQ Sequence 20 AA:

Query Match 30.4%; Score 31; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PEVEXK 20
| | | | |
Db 1 pevexk 6

RESULT 18
Y59576
ID Y59576 standard; peptide: 20 AA.
XX
AC Y59576;
XX
DT 03-APR-2000 (first entry)
XX
DE GAD65 fragment, peptide #16.
XX
GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
KW Insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
KM therapy.
XX
OS Homo sapiens.
XX
PN US5998366-A.
XX
PD 07-DEC-1999.
XX
PE 09-APR-1997; 97US-0827618.
XX
PR 07-JUN-1995; 95US-0485725.
PR 21-SEP-1990; 90US-0586536.
PR 18-JUN-1991; 91US-0716909.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Tobin AJ, Kaufman DL, Erlander MG;
XX
DR WPI; 2000-095930/08.
XX
PT Ameliorating glutamic acid decarboxylase associated autoimmune
PT disorders such as insulin dependent diabetes mellitus and stiff man
PT disease -
XX
PS Example 11; Column 42; 61pp; English.
XX
CC This sequence represents a fragment of the glutamic acid decarboxylase 65
CC (GAD65) protein. The invention relates to a method of ameliorating GAD
CC associated autoimmune disorders by administering a GAD65 peptide to the
CC patient. The method can be used for ameliorating GAD associated
CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
CC and stiff man disease. GAD65 can also be useful for screening drugs that
CC alter GAD function, for generating monoclonal antibodies and in
CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
CC and the diagnosis is quite easy. It is also possible to obtain much
CC larger quantities of polypeptide via recombinant techniques than are
CC available from natural sources.

XX
SQ Sequence 20 AA;

Query Match 30.4%; Score 31; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNMYAM 6
| | | | |
Db 15 snmyam 20

RESULT 19
Y59578
ID Y59578 standard; peptide: 20 AA.
XX
AC Y59578;
XX
DT 03-APR-2000 (first entry)
XX
DE GAD65 fragment, peptide #16.
XX
GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
KW Insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
KM therapy.
XX
OS Homo sapiens.
XX
PN US5998366-A.
XX
PD 07-DEC-1999.
XX
PE 09-APR-1997; 97US-0827618.
XX
PR 07-JUN-1995; 95US-0485725.
PR 21-SEP-1990; 90US-0586536.
PR 18-JUN-1991; 91US-0716909.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Tobin AJ, Kaufman DL, Erlander MG;
XX
DR WPI; 2000-095930/08.
XX
PT Ameliorating glutamic acid decarboxylase associated autoimmune
PT disorders such as insulin dependent diabetes mellitus and stiff man
PT disease -
XX
PS Example 11; Column 42; 61pp; English.
XX
CC This sequence represents a fragment of the glutamic acid decarboxylase 65
CC (GAD65) protein. The invention relates to a method of ameliorating GAD
CC associated autoimmune disorders by administering a GAD65 peptide to the
CC patient. The method can be used for ameliorating GAD associated
CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
CC and stiff man disease. GAD65 can also be useful for screening drugs that
CC alter GAD function, for generating monoclonal antibodies and in
CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
CC and the diagnosis is quite easy. It is also possible to obtain much
CC larger quantities of polypeptide via recombinant techniques than are
CC available from natural sources.

RESULT 20
 Y57065
 ID Y57065 standard; peptide; 20 AA.
 XX
 AC Y57065;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Glutamate decarboxylase peptide GAD p18.
 XX
 KW Neglected target tissue antigen; NTFA; autoimmunity; autoimmune response;
 KW immunotherapeutic agent; insulin dependent diabetes mellitus;
 KW multiple sclerosis; autoimmune thyroiditis; rheumatoid arthritis;
 KW uveoretinitis; inflammatory response.
 XX
 OS Homo sapiens.
 XX
 PN W09956763-A1.
 PD 11-NOV-1999.
 XX
 PF 07-MAY-1999; 99WO-US10250.
 XX
 PR 07-MAY-1998; 98US-0084636.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Kaufman DL, Tian J, Olcott A;
 XX
 DR WPI: 2000-052905/04.
 XX
 PT Administration of neglected target tissue antigens to modulate immune
 PT responses
 PS Disclosure: Page 23; 79pp; English.
 XX
 CC Amino acid sequences Y57063-Y57091 are examples of neglected target
 CC tissue antigens NTFA's. NTFA's are antigens (whole antigens or fragments)
 CC not involved in autoimmunity. These peptides and proteins are used in the
 CC method of the invention which involves administering an NTFA as an
 CC antigen based immunotherapeutic agent, to a host afflicted with an
 CC autoimmune response associated with an autoimmune disease. The
 CC immunotherapeutic agent is used to treat autoimmune diseases such as
 CC insulin dependent diabetes mellitus, multiple sclerosis, autoimmune
 CC thyroiditis, uveoretinitis, rheumatoid arthritis or abnormal inflammatory
 CC immune responses. The NTFA induces regulatory tolerance by elicitation of
 CC regulatory T cells among T cells recognizing the NTFA but not
 CC participating in the immune response. The NTFA are capable of recognition
 CC by substantial populations of uncommitted T cells which can be primed, or
 CC biased, towards regulatory responses to provide effective treatment. The
 CC NTFA are effective in regulating undesirable immune responses even when
 CC target determinants used as agents promoting tolerance agents have failed
 CC to induce an effective regulatory T cell response. NTFA's as agents
 CC promoting tolerance are anticipated to be safer than use of target
 CC determinants.
 CC
 XX
 SQ Sequence 20 AA;
 XX

Query Match: 30.4%; Score 31; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 PEVKEK 20
 |||||
 Db 1 pevkek 6

RESULT 21
 R43871
 ID R43871 standard; Protein; 14 AA.

XX
 AC R43871;
 XX
 DT 23-DEC-1993 (first entry)
 XX
 DE OmpA2-R-7 signal peptide.
 XX
 KW Polypeptide; antibodies; HTLV; AIDS; vaccine.
 KW
 XX
 OS Human T-cell lymphotropic virus.
 XX
 PN EP552850-A.
 PD 28-JUL-1993.
 XX
 PF 10-OCT-1985; 85EP-0307260.
 XX
 PR 10-OCT-1984; 84US-0659339.
 PR 23-JAN-1985; 85US-0693866.
 XX
 PA (CENZ) CENTOCOR INC.
 XX
 PI Chang NT, Gallo RC, Wong-staal F;
 XX
 DR WPI: 1993-236543/30.
 DR N-PSDB: Q45924.
 XX
 PT Cloning and expression of new HTLV-III DNA - used to obtain
 PT polypeptide(s) and antibodies for diagnosis, prevention and
 PT treatment of HTLV-III infection, partic. AIDS
 XX
 PS Disclosure: Figure 6a; 31pp; English.
 XX
 CC A fragment of DNA approximately 200-500 base pairs in length is
 CC ligated into a recombinant vector (ompA1-R-6; ompA2-R-7 or ompA3-R-3)
 CC and used to transform E.coli. These cells then express a polypeptide
 CC which is immunoreactive with HTLV-III-specific antibody. The
 CC HTLV-III polypeptides can be used for the production
 CC of antibodies, in immunoassays for the detection of HTLV-
 CC III-specific antibodies and in vaccines for the prevention of AIDS.
 CC The antibodies can also be used to detect HTLV-III polypeptides.
 XX
 SQ Sequence 14 AA;
 XX

Query Match: 29.4%; Score 30; DB 14; Length 14;
 Best Local Similarity 54.5%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 9 ARFKMEPEVKE 19
 | | | | |
 Db 4 aeftpikvke 14

RESULT 22
 R96007
 ID R96007 standard; peptide; 15 AA.
 XX
 AC R96007;
 XX
 DT 19-FEB-1997 (first entry)
 XX
 DE Nuclear localisation signal.
 XX
 KW Target binding assembly; nucleic acid recognition unit; NF-kappa-B; SPI;
 KW TATA; human papillomavirus; HPV; HPV E2; human immunodeficiency virus;
 KW HIV LTR; Tat binding unit; probe; assembly sequence; asymmetry sequence;
 KW nuclear localisation signal sequence; human cell; HPV LTR; HIV; bacteria;
 KW virus.
 XX
 OS Synthetic.
 XX
 PN W09617956-A2.

XX 13-JUN-1996.
 XX
 PD 07-DEC-1995; 95WO-US15944.
 PF
 XX 09-DEC-1994; 94US-0353476.
 PR
 XX (GENE-) GENE POOL INC.
 PA
 XX Weininger AM, Weininger S;
 PI
 XX WPI; 1996-287199/29.
 DR
 XX Probe nucleic acids, target binding assemblies, etc - for detection
 PT and localisation of specific nucleic acid sequences, esp. HIV and
 PT HPV
 PS
 XX Claim 23; Page 114; 172pp; English.
 CC This sequence represents a nuclear localisation sequence of a target
 CC binding assembly (TBA) of the invention. The TBA is recognised by the
 CC target binding region (TBR) of a probe of the invention. The probe of
 CC the invention contains a TBR, a booster binding region (BBR), and an
 CC optional support or attachment (OSA). The TBA contains at least one
 CC nucleic acid recognition unit (NAR), and optionally a linker sequence, an
 CC assembly sequence, an asymmetry sequence, a nuclear localisation signal
 CC sequence, and an OSA. The assembly sequence and asymmetry sequences are
 CC responsible for the folding and association of the NARs. The NARs are
 CC selected from NF-kappa-B, SPI, TARA, human papillomavirus (HPV) E2, HPV
 CC LTR, human immunodeficiency virus (HIV) LTR and Tat binding units. The
 CC linker sequence is an oligopeptide, which does not interfere with NAR
 CC function, but provides stability and control over the spacing of the NAR
 CC from the rest of the TBA. The OSA is an attached support or indicator,
 CC or other means of localisation of the probe. The probe can be used in a
 CC method for detecting or localising a specific target nucleic acid
 CC sequence (TNA). The method is highly sensitive, and has a high degree of
 CC specificity. The method can be used for detecting specific nucleic acid
 CC sequences, including those found in human cells, in HIV, HPV, and other
 CC nucleic acid containing systems, including bacteria and viruses.
 CC
 SQ Sequence 15 AA;
 XX
 XX
 Query Match 29.4%; Score 30; DB 17; Length 15;
 Best Local Similarity 57.1%; Pred. No. 34;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 14 FPEYKER 20
 DB 2 ypeikdk 8
 XX
 RESULT 23
 ID R06363
 XX R06363 standard; protein; 22 AA.
 AC R06363;
 XX
 DT 13-DEC-1990 (first entry)
 XX
 DE C-terminal of fish growth hormone.
 XX
 KW Osteichthyes; flatfish.
 KM
 XX Pleuronectiformes paralicthys.
 OS
 XX JP02174799-A.
 PN
 XX 06-JUL-1990.
 PD
 XX 20-OCT-1988; 88JP-0264928.
 PF
 XX 26-SEP-1988; 88JP-0240265.
 PR

PR 26-AUG-1988; 88JP-0211694.
 XX
 XX (NIOC) NIPPON OIL KK.
 PA
 XX WPI; 1990-250742/33.
 DR
 XX Fish growth hormone - extracted from pituitary gland of
 PT Pleuronectiformes Paralicthys fish.
 PT
 XX Claim 3; Page 1; 6pp; English.
 PS
 XX The protein is extracted from the pituitary gland of P. paralicth-
 CC thys. It can be used to accelerate the growth of the species of
 CC Pleuronectiformes and Salmo oncorhynchus. The protein has a mol.
 CC wt. of 22,000 and an isoelectric pt. of 7.1.
 CC See also R06362, R07480 and R07481.
 CC
 SQ Sequence 22 AA;
 XX
 XX
 Query Match 29.4%; Score 30; DB 11; Length 22;
 Best Local Similarity 36.4%; Pred. No. 51;
 Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 6 MMARFRKPE 16
 DB 7 ltvakrllpe 17
 XX
 RESULT 24
 ID W63838
 XX W63838 standard; Protein; 25 AA.
 AC W63838;
 XX
 DT 01-OCT-1998 (first entry)
 XX
 DE Plasmid PRMP100 M.HpyI protein fragment from H. pylori strain 60190.
 XX
 KW IcaA; immunoassay; detection; ulcerogenic; gastric carcinoma; treatment;
 KW peptic ulcer; immunisation; vaccine; protection.
 KM
 XX Synthetic.
 OS Helicobacter pylori.
 XX
 PN WO9743901-A1.
 XX
 PD 27-NOV-1997.
 XX
 PF 20-MAY-1997; 97WO-US08558.
 XX
 PR 20-MAY-1996; 96US-0650528.
 XX
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Blaser MJ, Miller GG, Peek RM, Thompson SA;
 XX
 DR WPI; 1998-286350/25.
 DR N-PSDB; V43963.
 XX
 PT New Helicobacter pylori proteins - induced by contact with
 PT epithelium and related DNA, are associated with ulcer formation,
 PT useful in diagnosis and immunisation
 XX
 XX Disclosure; Fig 2; 107pp; English.
 PS
 XX This sequence is a M.hpyI protein fragment found in plasmid PRMP100. This
 CC plasmid also contains fragments of proteins CysE and IcaA derived from
 CC Helicobacter pylori strain 60190. This plasmid is used in method which
 CC provides an IcaA protein or fragments of the protein capable of being
 CC used in standard immunoassays to detect H. pylori-specific antibodies,
 CC particularly for diagnosis, especially antibodies characteristic of
 CC IcaA-positive strains which are ulcerogenic. Detecting presence of

CC Igea-positive strains also allows the risk of developing gastric
CC carcinoma to be assessed. Ligands, particularly antibodies, that
CC recognise Igea proteins are used to treat peptic ulcers, while
CC immunisation with Igea-negative H₂ pylori is used to protect against
CC infection (and its consequences such as ulcers, gastritis and gastric
CC cancer). Immunogenic Igea fragments, or the nucleic acid encoding them,
CC can also be used for vaccination. Antibodies (Ab) raised against Igea can
CC be used therapeutically or to screen other strains for homologous
CC proteins. Expression of Igea is strongly correlated with ulceration, so
CC detecting Igea allows differentiation between ulcerogenic and
CC non-ulcerogenic strains.
XX
50 Sequence 25 AA;

CC cancer) Immunogenic IGEA fragments, or the nucleic acid encoding them,
 CC can also be used for vaccination. Antibodies (Ab) raised against IGEA can
 CC be used therapeutically or to screen other strains for homologous
 CC proteins. Expression of IGEA is strongly correlated with ulceration, so
 CC detecting IGEA allows differentiation between ulcerogenic and
 CC non-ulcerogenic strains.
 XX
 SQ Sequence 25 AA;

 Query Match 29.4%, Score 30, DB 19, Length 25;
 Best Local Similarity 36.4%, Pred. No. 58;
 Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0.

Query Match	29.4%	Score 30;	DB 19;	Length 25;
Best Local Similarity	36.4%	Pred. No. 58;		
Matches	4; conservative	5; mismatches	2; indels	0; gaps
OY	9 ARFKMPPEYKE 19			
	..::: : :			
Db	6 skykliprike 16			

Query Match	29.4%	Score 30;	DB 19;	Length 25;
Best Local Similarity	36.4%	Pred. No. 58;		
Matches	4;	Conservative	5;	Mismatches
			2;	Indels
				0;
Gaps				0;
Oy	9	ARFKMPEYKE	19	
	:::: :	::		
Db	6	skYkIPIfike	16	

Search completed: February 5, 2001, 10:47:30
Job time: 624 sec

RESULT	25	
W63841		
ID	W63841	standard; Protein; 25 AA.
XX		
AC	W63841;	
XX		
DT	01-OCT-1998	(first entry)
XX		
DE	Plasmid pMP100 M.HpyI protein fragment from H. pylori strain J178.	
XX		
KW	Icga; immunoassay; detection; ulcerogenic; gastric carcinoma; treatment;	
RW	peptic ulcer; immunisation; vaccine; protection.	
XX		
OS	Synthetic.	
XX	Helicobacter pylori.	
PN	W09743901-A1.	
XX		
PD	27-NOV-1997.	
XX		
PF	20-MAY-1997; 97WO-US08558.	
XX		
PR	20-MAY-1996; 96US-0650528.	
XX		
PA	(UTVA-) UNIV VANDERBILT.	
XX		
PI	Blaser MJ, Miller GG, Peek RM, Thompson SA;	
XX		
DR	WPI; 1998-286350/25.	
XX		
DR	N-PSDB; VA4037.	
XX		
PT	New Helicobacter pylori proteins - induced by contact with	
PT	epithelium and related DNA, are associated with ulcer formation,	
XX	useful in diagnosis and immunisation	
PS	Disclosure; Fig 2; 107pp; English.	
XX		
CC	This sequence is a M.hpyI protein fragment found in plasmid pMP100. This	
CC	plasmid also contains fragments of proteins CysE and IcaA derived from	
CC	Helicobacter pylori strain J178. This plasmid is used in method which	
CC	provides an IcaA protein or fragments of the protein capable of being	
CC	used in standard immunoassays to detect H. pylori-specific antibodies,	
CC	particularly for diagnosis, especially antibodies characteristic of	
CC	IcaA-positive strains which are ulcerogenic. Detecting presence of	
CC	IcaA-positive strains also allows the risk of developing gastric	
CC	carcinoma to be assessed. Ligands, particularly antibodies, that	
CC	recognise IcaA proteins are used to treat peptic ulcers, while	
CC	immunisation with IcaA-negative H. pylori is used to protect against	
CC	infection (and its consequences such as ulcers, gastritis and gastric	

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:47:30 ; Search time 68.99 Seconds
(without alignments)
9.913 Million cell updates/sec

Title: US-08-981-824-3

Perfect score: 113

Sequence: 1 NWELADQPQNLIELMHGQT 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_36:*

1: /SIDSL/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	113	100.0	20	W18844	65 kD Glutamic ac
2	113	100.0	20	W01795	Human 65 kD glutam
3	86	76.1	22	R72270	Glutamic acid deca
4	86	76.1	22	R72271	GAD65 fragment, pe
5	50	44.2	20	R72271	Glutamic acid deca
6	50	44.2	20	Y59551	GAD65 fragment, pe
7	35	31.0	15	W63553	Beta (1 -> 4)-N-ac
8	35	31.0	21	W31018	Mugwort pollen all
9	34	30.1	23	W31051	Mugwort pollen all
10	33	29.2	17	Y14278	Chemokine peptide,
11	33	29.2	23	R20273	Beta-2 integrin pe
12	33	29.2	23	W02061	Human beta2 integr

13	33	29.2	25	17	R99561
14	33	29.2	25	17	R92504
15	33	29.2	25	18	W15702
16	33	29.2	25	18	W24260
17	33	29.2	25	21	Y53906
18	33	28.3	17	20	Y34179
19	32	28.3	20	20	Y42241
20	32	28.3	20	20	Y43409
21	32	28.3	22	20	Y42236
22	32	28.3	22	20	Y42240
23	32	28.3	22	20	Y43404
24	32	28.3	22	20	Y43408
25	32	28.3	23	20	Y42239
26	32	28.3	23	20	Y43407
27	31	27.4	12	20	Y24751
28	31	27.4	20	16	R74066
29	31	27.4	20	20	Y39986
30	31	27.4	20	20	W89562
31	31	27.4	25	20	W97645
32	30	26.5	13	12	R12838
33	30	26.5	18	9	P82907
34	29	25.7	14	16	W21324
35	29	25.7	14	19	Y20370
36	29	25.7	15	16	R93652
37	29	25.7	15	18	W17995
38	29	25.7	16	16	R82544
39	29	25.7	17	20	Y41871
40	29	25.7	18	17	W05200
41	29	25.7	20	18	W16845
42	29	25.7	20	19	W40241
43	29	25.7	22	12	R1415
44	29	25.7	25	17	R85706
45	28.5	25.2	21	18	W16211

ALIGNMENTS

RESULT	1	
W18844	standard; peptide; 20 AA.	
W18844:		
05-JAN-1998 (first entry)		
65 kD Glutamic acid decarboxylase peptide fragment III.		
GAD: 65 kD: human; glutamic acid decarboxylase; autoreactive; diagnosis;		
insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;		
predisposition; autoimmune; tumour; rheumatoid arthritis;		
multiple sclerosis.		
Synthetic.		
DE19526561-AA.		
23-JAN-1997.		
20-JUL-1995; 95DE-1026561.		
20-JUL-1995; 95DE-1026561.		
(BOEF) BOEHRINGER MANNHEIM GMBH.		
Donle F, Endl J, Ganz M, Jung G, Klentech-engel R;		
Pozzilli P, Stahl P;		
WPI: 1997-088254/09.		
Skin test for diagnosis of cell-mediated diseases, esp. diabetes -		
Involving intradermal admn. of auto-reactive substances		

Interferon-gamma 1
Interferon gamma p
Internal fragment
Peptide #2 of huma
Fragment derived f
Human pre-prohepar
Multivalent ligand
Multivalent ligand
Multivalent ligand
Multivalent ligand
Multivalent ligand
Multivalent ligand
Multivalent ligand
Lactoferrin duodec
Superantigen pepit
FIV ORF(51-71) pep
Recombinant heat-r
Breast-specific BS
HIV protecting vac
Acetylcholinestera
Glucagon precursor
Human microtubule
HIV principal neut
Helicobacter pylor
G. oxydans T100 L-
Rheumatoid arthrit
Helicobacter fam1
Helicobacter pylor
H. pylori Urease A
Human immunodefici
Human cytokeratin
Peptide containing

```

Ps Claim 11; Page 9; 12pp; German.
Cc W18842-70 are peptide fragments of the 65 kD human glutamic acid
Cc decarboxylase (GAD). The fragments are autoreactive substances used for
Cc diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
Cc determined by using a claimed method for diagnosis of cell-mediated
Cc diseases or a predisposition to cell-mediated diseases, which is effected
Cc by administering an autoreactive substance intradermally and establishing
Cc the diagnosis on the basis of the occurrence or lack of a positive
Cc reaction at the site of administration. The method is used for diagnosis
Cc of autoimmune and tumour diseases, preferably T-cell-mediated diseases
Cc such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
Sx
SQ Sequence 20 AA:
Query Match 100.0%; Score 113; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 NWEIADQFONLEIEILHMCOT 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 nweiadqpnleielmhqt 20
RESULT 2
W01795 W01795 standard; peptide; 20 AA.
AC W01795:
XX 15-OCT-1997 (first entry)
XX DT
XX Human 65 kD glutamine decarboxylase peptide.
XX DE
XX Humam; glutamine decarboxylase; GAD; diagnosis; predisposition;
XX KW tumour; immunologicall; disease; autoimmune; diabetes; reagent;
XX KM determination; T cell; subpopulation; medicament; treatment;
XX KW prevention; production; antigen; immunogen; tolerogen; isolation;
XX reinjection; inactivation.
XX OS
XX Homo sapiens.
XX XX
XX DEL9525784-A1.
XX PN
XX 16-JAN-1997.
XX PD
XX PF 14-JUL-1995; 95DE-1025784.
XX XX
XX PR 14-JUL-1995; 95DE-1025784.
XX XX
XX (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX PA
XX Albert W, Boitard C, Endl J, Jung G, Schendel D;
XX PI Stahl P, Van Enderst P;
XX PI
XX MPI; 1997-078452/08.
XX DR
XX Glutamine decarboxylase peptide(s) - for diagnosis and therapy of
XX PT diabetes, etc.
XX PS
XX Claim 1; Page 12; 15pp; German.
Cc The present peptide is a fragment of the human 65 kD glutamine
Cc decarboxylase (GAD), which can be used to diagnose, or diagnose a
Cc predisposition to, a tumour or immunological disease, preferably an
Cc autoimmune disease, especially diabetes. It can also be used as a
Cc reagent to determine specific T cell subpopulations, in medicaments
Cc to treat or prevent immunological diseases, preferably autoimmune
Cc diseases, especially diabetes, to produce antigens, especially
Cc immunogens or tolerogens and to isolate specific T cell
Cc subpopulations, which can be used to produce antigens or for
Cc reinjection, optionally after inactivation.

```

```

xx
SQ      Sequence      20 AA:

Query Match      100.0%; Score 113; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 6,9e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 NWELADQPONLEELIMHCOT 20
        |||
Db       1 nweledqpnlleelimhcot 20

RESULT      3
R72270
ID      R72270 standard; Peptide; 22 AA.
XX
AC      R72270;
XX
DT      13-NOV-1995 (first entry)
XX
DE      Glutamic acid decarboxylase (GAD65) fragment.
XX
KW      Glutamic acid decarboxylase; GAD65; autoimmune disorders;
KW      Insulin-dependent diabetes mellitus; stiff man disease.
XX
OS      Homo sapiens.
XX
PN      W09507992-A.
XX
PD      23-MAR-1995.
XX
PE      24-AUG-1994; 94WO-US09478.
XX
PR      17-SEP-1993; 93US-0123859.
XX
PA      (REGC ) UNIV CALIFORNIA.
PI      Clare-Salzer MJ, Erlander MG, Kaufman DL, Tobin AJ;
PI      WPI; 1995-131360/17.
XX
DR      New polypeptide fragments of glutamic acid decarboxylase - for
PT      diagnosis and treatment of autoimmune disease, esp. insulin
PT      dependent diabetes; also related nucleic acid, vectors,
PT      antibodies, hybridoma(s) etc.
XX
PS      Example 11: Page 76; 100pp; English.
XX
CC      Q86481 and Q86482 encode R7173 and R79105, rat and human glutamic
CC      acid decarboxylase (GAD65) respectively, from which the GAD65
CC      fragments described in R72261-R72298 were derived. These fragments
CC      can be used to detect autoantibodies against GAD, e.g. to diagnose
CC      and treat GAD-related autoimmune disorders, such as insulin
CC      dependant diabetes mellitus or stiff man disease.
XX
SQ      Sequence      22 AA:

Query Match      76.1%; Score 86; DB 16; Length 22;
Best Local Similarity 100.0%; Pred. No. 1,8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 NWELADQPONLEELIM 16
        |||
Db       7 nweledqpnlleelim 22

RESULT      4
Y59571
ID      Y59571 standard; peptide; 22 AA.
XX
AC      Y59571;
XX

```

[illegible]

```
XX      23-MAR-1995 .
PD
XX
PF      24-AUG-1994;       94WO-US09478.
XA
PR      17-SEP-1993;       93US-0123859.
XX
PA      (REGC ) UNIV CALIFORNIA.
XX
PI      Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
DR      WPI: 1995-131360/17.
XX
PT      New polypeptide fragments of glutamic acid decarboxylase - for
PR      diagnosis and treatment of autoimmune disease, esp. insulin
PT      dependent diabetes, also related nucleic acid, vectors,
PP      antibodies, hybridoma(s) etc.
XX
PS      Claim 1; Page 76; 100pp; English.
XX
CC      Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
CC      acid decarboxylase (GAD65) respectively, from which the GAD65
CC      fragments described in R72261-R72298 were derived. These fragments
CC      can be used to detect autoantibodies against GAD, e.g. to diagnose
CC      and treat GAD-related autoimmune disorders, such as Insulin
CC      dependant diabetes mellitus or stiff man disease.
XX
SQ      Sequence    20 AA:

Query Match          44.2%; Score 50; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches   9; Conservative   0; Mismatches     0; Indels     0; Gaps     0.

OY      12 EELIMHCQT 20
        |||||||
Db       1 eelimhcqt 9

RESULT      6
ID         YS9551
AC         YS9551: standard; peptide; 20 AA.
XX
AC         YS9551:
XX
DT         03-APR-2000 (first entry)
XX
DE         GAD65 fragment, peptide #11.
XX
KW         GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
RM         insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
RW         therapy.
XX
OS         Homo sapiens:
XX
PN         US5998366-A.
XX
PD         07-DEC-1999.
XX
PE         09-APR-1997;       97US-0827618.
XX
PR         07-JUN-1995;       95US-0485725.
PR         21-SEP-1990;       90US-0586536.
PR         18-JUN-1991;       91US-0716909.
XX
PA         (REGC ) UNIV CALIFORNIA.
XX
PI         Tobin AJ, Kaufman DL, Erlander MG;
DR         WPI: 2000-095930/08.
XX
PT         Ameliorating glutamic acid decarboxylase associated autoimmune
PT         disorders such as insulin dependent diabetes mellitus and Stiff man
```

PT disease -
 XX Claim 1; Column 42; 61pp; English.
 PS
 XX
 CC This sequence represents a fragment of the glutamic acid decarboxylase 65
 CC (GAD65) protein. The invention relates to a method of ameliorating GAD
 CC associated autoimmune disorders by administering a GAD65 peptide to the
 CC patient. The method can be used for ameliorating GAD associated
 CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
 CC and stiff man disease. GAD65 can also be useful for screening drugs that
 CC alter GAD function, for generating monoclonal antibodies and in
 CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
 CC and the diagnosis is quite easy. It is also possible to obtain much
 CC larger quantities of polypeptide via recombinant techniques than are
 CC available from natural sources.
 CC
 SQ Sequence 20 AA;
 Query Match 44.2%; Score 50; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 EETLMHCOT 20
 Db 1 eettlmhct 9
 RESULT 7
 W63553
 ID W63553 standard; peptide: 15 AA.
 XX
 AC W63553;
 XX
 DT 24-NOV-1998 (first entry)
 XX
 DE Beta (1 -> 4)-N-acetylglucosaminyl-transferase (Gnt-IV) fragment.
 XX
 KW Beta (1 -> 4)-N-acetylglucosaminyl-transferase; Gnt-IV; bovine; human;
 KW enzyme; sugar chain subunit; branched oligosaccharide; polysaccharide;
 KW drug; reagent; food; biopolymer; glycoprotein; erythropoietin.
 XX
 OS Homo sapiens.
 OS Bos sp.
 XX
 PN W09826053-A1.
 XX
 PD 18-JUN-1998.
 XX
 PF 10-DEC-1997; 97WO-JP04546.
 XX
 PR 18-JUN-1997; 97JP-0161462.
 PR 12-DEC-1996; 96JP-0332411.
 XX
 PA (KIRI) KIRIN BEER KK.
 XX
 PI Minowa M, Oguri S, Takeuchi M, Taniguchi N, Yoshida A;
 XX
 DR WPI: J998-348516/30.
 XX
 PT Recombinant beta(1-4)-N-acetylglucosaminyl-transferase - allows
 PT production of difficultly accessible branched poly:saccharides for
 PT food and drug use
 XX
 PS Claim 4; Page 49; 112pp; Japanese.
 XX
 CC Sequences shown in W63543 to W63556 represent fragments of beta (1 -> 4)-
 CC N-acetylglucosaminyl-transferase (Gnt-IV) enzyme. The invention provides
 CC bovine and human Gnt-IV enzymes that can be used for converting sugar
 CC chain subunits having one structure to another structure. Vectors
 CC containing the DNA sequences encoding these enzymes can be used to
 CC transform host cells for the production of the Gnt-IV enzymes. The
 CC enzymes are useful in the production of branched oligosaccharides and

CC polysaccharides which are difficult of access by other methods. They are
 CC also useful in the production of drugs, reagents and foods and in
 CC modifying the properties of biopolymers containing sugar chains. The
 CC enzyme may also be used for the preparation of glycoproteins such as
 CC erythropoietin.
 CC
 SQ Sequence 15 AA;
 Query Match 31.0%; Score 35; DB 19; Length 15;
 Best Local Similarity 50.0%; Pred. No. 19;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 6 DQPNDSEETLMH 17
 Db 1 dkpvnvessylfh 12
 RESULT 8
 W31018
 ID W31018 standard; peptide: 21 AA.
 XX
 AC W31018;
 XX
 DT 09-JAN-1998 (first entry)
 XX
 DE Mugwort pollen allergen B cell epitope.
 XX
 KW Cofactor-independent phosphoglycerate mutase; PGM-1; E.C. 5.4.21;
 KW Timothy grass; pollen; allergy; plant allergen; panallergen; B cell;
 KW T cell; epitope; immunotherapy; detection; diagnosis; hay fever;
 KW conserved.
 XX
 OS Artemisia vulgaris.
 OS
 PN W09705258-A2.
 XX
 PD 13-FEB-1997.
 XX
 PF 02-AUG-1996; 96WO-AT00141.
 XX
 PR 02-AUG-1995; 95AT-0001320.
 XX
 PA (BIOM) BIOMAY PRODN & HANDELS GMBH.
 XX
 PI Breitenbach M, Ebner C, Engel E, Ferreira F, Jilek A;
 PI Kraft D, Richter K, Rheinberger H;
 XX
 DR WPI: 1997-145695/13.
 XX
 PT New recombinant DNA encoding plant phosphoglycerate mutase or its
 PT antigenic epitope(s) - useful for diagnosis or treatment of
 PT allergies to pollen and plant-derived foods
 XX
 PS Disclosure; Fig 11a; 160pp; German.
 XX
 CC W31018-W31040 are B cell epitopes of mugwort pollen co-factor-
 CC independent phosphoglycerate mutase (PGM-1) isoform Art6. PGM-1 is
 CC a highly conserved plant allergen (panallergen) which can cause
 CC cross-reactivity in patients allergic to pollen and plant-derived
 CC foods. PGM-1 and it's B cell and T cell epitopes can be used for the
 CC in vitro detection of allergy against PGM-1, by measuring serum IgE
 CC or a cellular reaction. They can also be used in immunotherapy and
 CC will not cause an autoimmune response because PGM-1 is significantly
 CC different from the human enzyme, which is co-factor dependent.
 CC
 SQ Sequence 21 AA;
 Query Match 31.0%; Score 35; DB 18; Length 21;
 Best Local Similarity 55.6%; Pred. No. 28;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 NWEADPOQ 9
 DB 7 swkladphk 15

RESULT 9
 W31051
 ID W31051 standard; peptide: 23 AA.
 AC W31051;
 XX
 DT 09-JAN-1998 (first entry)
 DE Mugwort pollen allergen B cell epitope.
 XX
 XX Cofactor-independent phosphoglycerate mutase; PGM-1; E.C. 5.4.21;
 KM Timothy grass; pollen; allergy; plant allergen; panallergen; B cell;
 KM T cell; epitope; immunotherapy; detection; diagnosis; hay fever;
 KM conserved.
 XX Artemisia vulgaris.
 XX WO9705258-A2.
 XX
 XX 13-FEB-1997.
 PD
 XX 02-AUG-1996; 96WO-AT00141.
 PF
 XX 02-AUG-1995; 95AT-0001320.
 PR
 XX (BIOM-) BIOMAX PRODN & HANDELS GMBH.
 PA
 PI Breitenbach M, Ebner C, Engel E, Ferreira F, Jilek A;
 PI Kraft D, Richter K, Rheinberger H;
 PI
 PI MPI: 1997-145695/13.
 DR
 XX
 XX New recombinant DNA encoding plant phosphoglycerate mutase or its
 PT antigenic epitope(s) - useful for diagnosis or treatment of
 PT allergies to pollen and plant-derived foods
 PS
 PS Disclosure: Fig 11b; 160pp; German.
 XX
 CC W31051-W31072 are B cell epitopes of mugwort pollen co-factor-
 CC independent phosphoglycerate mutase (PGM-1) isoform Art17. PGM-1 is
 CC a highly conserved plant allergen (panallergen) which can cause
 CC cross-reactivity in patients allergic to pollen and plant-derived
 CC foods. PGM-1 and its B cell and T cell epitopes can be used for the
 CC in vitro detection of allergy against PGM-1, by measuring serum IgE
 CC or a cellular reaction. They can also be used in immunotherapy and
 CC will not cause an autoimmune response because PGM-1 is significantly
 CC different from the human enzyme, which is co-factor dependent.
 CC
 SQ Sequence 23 AA;

Query Match 30.1%; Score 34; DB 18; Length 23;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 WELADPOQ 9
 DB 10 wkldphk 17

RESULT 10
 ID Y14278 standard; peptide: 17 AA.
 AC Y14278;
 XX
 XX 29-JUL-1999 (first entry)
 DT
 XX

DE Chemokine peptide, CRD CysOCys16peptide(1-15)[SDF1].
 XX
 XX Chemokine; immune response; monocyte chemoattractant protein-1; MCP-1;
 KM chemokine-induced activity; inflammatory response; vascular indication;
 KM haematopoietic cell-associated activity; tumour; coronary artery disease;
 KM myocardial infarction; unstable angina pectoris; atherosclerosis; asthma;
 KM vasculitis; lentiviral infection; low bone mineral density; suppressor;
 KM parasitic infection; autoimmune disease; psoriasis; wound healing;
 KM organ transplant rejection; rheumatoid arthritis; allergy; therapy;
 KM arachidonic acid pathway.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 PN WO9912968-A2.
 XX
 XX 18-MAR-1999.
 PD
 XX 11-SEP-1998; 98WO-US19052.
 PF
 XX 11-SEP-1997; 97US-0927939.
 PR
 XX (NEOR-) NEORX CORP.
 PA
 PI Gralinger DJ, Kanaly ST, Tatalick LM;
 PI MPI: 1999-347124/29.
 DR
 XX
 XX New chemokine peptides and mimetics
 PT
 PT
 PS Disclosure: Fig 14; 208pp; English.
 XX
 XX This sequence represents a fragment of the chemokine SDF1.
 CC The invention relates to chemokine peptides and mimetics, particularly
 CC derived from monocyte chemoattractant protein-1 (MCP-1). The chemokine
 CC peptides and variants and derivatives can inhibit or reduce or increase,
 CC or enhance chemokine-induced activity. They can be used for increasing or
 CC enhancing an inflammatory response, an immune response or haematopoietic
 CC cell-associated activity at a tumour site. They can also be used for
 CC preventing or inhibiting an indication associated with haematopoietic
 CC cell recruitment or histamine release from basophils or mast cells. They
 CC can also be used to modulate the chemokine-induced activity of
 CC haematopoietic cells at a preselected physiological site, to treat a
 CC vascular indication, e.g. coronary artery disease, myocardial infarction,
 CC unstable angina pectoris, atherosclerosis, or vasculitis, lentiviral
 CC infection or replication (e.g. HIV), low bone mineral density, a
 CC parasitic infection in a vertebrate animal (e.g. malaria), an autoimmune
 CC disease, to suppress tumour growth in a vertebrate animal, to prevent or
 CC treat psoriasis in a mammal, to enhance wound healing, to prevent or
 CC treat asthma, organ transplant rejection, rheumatoid arthritis or
 CC allergy. They can also be used to inhibit a product or intermediate in
 CC the arachidonic acid pathway and where leukotriene, thromboxane and/or
 CC prostaglandin are inhibited and to prevent or inhibit an indication
 CC associated with elevated TNF-alpha.
 CC
 SQ Sequence 17 AA;

Query Match 29.2%; Score 33; DB 20; Length 17;
 Best Local Similarity 50.0%; Pred. No. 46;
 Matches 8; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

OY 3 ELADPOQNEEILMHC 18
 DB 4 glacnptnl--lklhc 17

RESULT 11
 ID R20273 standard; peptide: 23 AA.
 AC R20273;
 XX
 XX

DT 10-APR-1992 (first entry)

XX Beta-2 integrin peptide derived from CD11b alpha subunit.

DE Immune response; phagocyte-mediated tissue damage; A domain; MM;

XX Inflammation; CD11/CD18 complex; alpha subunit; CR3; M01; Mav-1;

KW complement receptor type 3; heterodimer.

XX Synthetic.

OS

XX W09119511-A.

XX

XX 26-DEC-1991.

PD

XX 18-JUN-1991; 91WO-US04338.

PF

XX 04-JAN-1991; 91US-0637830.

PR

XX 18-JUN-1990; 90US-0539842.

PR

XX (GEHO-) GEN HOSPITAL CORP.

PA

XX Arnaout MA;

PI

XX WPI; 1992-024197/03.

DR

XX Beta-2 integrin peptide CD11b, recombinant hetero-dimer

PT CD11b/CD18 - or MAb against them; useful for inhibiting CD11/CD18

PT mediated immune response in control of phagocyte-mediated tissue

PT damage

XX Claim 8; Page 73; 84pp; English.

PS

XX The peptide (SEQ ID NO: 20) corresponds to residues 112-134 of

CC CD11b alpha subunit of beta 2 integrin. The sequence occurs

CC immediately prior to the A domain which is conserved in the CD11c

CC and CD1a subunits and also in the collagen/heparin/platelet GRI

CC binding regions of the mature von Willebrand factor. The peptide

CC may be synthesised or prepd. by recombinant techniques using the

CC gene, the sequence of which was disclosed by Arnaout et al., J.

CC Cell Biol. 106:2153 (1988). (References are also provided for the

CC DNA sequences of human CD18, CD1c and CD1a). The peptide is cap-

CC able of inhibiting a CD11/CD18 mediated immune response and is

CC useful for treatment of ischaemia reperfusion injury, burns,

CC frostbite, acute arthritis, asthma and adult respiratory distress

CC syndrome. It may also be used to block intra-islet infiltration of

CC macrophages associated with insulin-dependent diabetes mellitus,

CC and for controlling phagocyte-mediated tissue damage to heart

CC muscle during acute cardiac insufficiency.

CC See also R20256-R20299.

CC

XX Sequence 23 AA;

SO

Query Match 29.2%; Score 33; DB 13; Length 23;

Best Local Similarity 46.7%; Pred. No. 66;

Matches 7; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 LADPQNLEILMHC 18

DB 3 lrgpqkfpalrgc 17

RESULT 12

W02061

ID W02061 standard; peptide; 23 AA.

XX

XX W02061;

AC

XX 09-APR-1997 (first entry)

DT

XX Human beta2 integrin subunit CD11b, residues 112-134.

DE

XX Beta2 integrin; A-domain; metal binding domain; inflammatory response;

XX

KW Immune response; inhibition; phagocyte-mediated tissue injury;

KW Inflammation.

XX

XX Homo sapiens.

OS

XX W09624063-A1.

XX

XX 08-AUG-1996.

PD

XX 30-JAN-1996; 96WO-US01314.

PF

XX 30-JAN-1995; 95US-0380167.

PR

XX (GEHO) GEN HOSPITAL CORP.

PA

XX Arnaout MA;

PI

XX WPI; 1996-371576/37.

DR

XX In vitro identification of integrin function antagonists - by

PT measuring binding of A-domain peptide derived from integrin to

PT ligand in presence and absence of candidate antagonist

PT

XX Disclosure; Page 11; 11pp; English.

PS

XX The sequences given in W02037-80 represent peptides derived from

CC beta2 integrin, esp. A-domain and the metal binding domains. These

CC peptides were selected using the method of the invention which

CC screens compounds for their ability to inhibit the binding of a

CC selected integrin to a ligand which naturally binds to it. The

CC method comprises measuring the binding of an A-domain peptide

CC derived from the selected integrin, to the ligand in the presence

CC and absence of the test compound and determining whether the binding

CC is decreased. Identified compounds are capable of interfering with

CC certain cellular immune/inflammatory responses, particularly

CC phagocyte-mediated tissue injury and inflammation. The numbering

CC of the amino acid residues is based on the deduced amino acid

CC sequence of the open reading frame of human CD11b from Arnaout et al.,

CC J. Cell. Biol. 106:2153 (1988).

CC

XX Sequence 23 AA;

SO

Query Match 29.2%; Score 33; DB 17; Length 23;

Best Local Similarity 46.7%; Pred. No. 66;

Matches 7; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 LADPQNLEILMHC 18

DB 3 lrgpqkfpalrgc 17

RESULT 13

R99561

ID R99561 standard; Peptide; 25 AA.

XX

XX R99561;

AC

XX 29-SEP-1996 (first entry)

DT

XX Interferon-gamma inducer protein tryptic peptide (aa79-103).

DE

XX Interferon-gamma inducer protein; IFN-gamma; antiviral; virucide;

KW antitumour; antibacterial; immunoregulator; adoptive immunotherapy;

KW therapy; cancer.

XX

XX Mus sp.

OS

XX EP12931-A2.

PN

XX 22-MAY-1996.

PD

XX 10-NOV-1995; 95EP-0308055.

PF

XX 29-SEP-1995; 95JP-0274988.
 PR 15-NOV-1994; 94JP-0304203.
 PR 23-FEB-1995; 95JP-0058240.
 PR 10-MAR-1995; 95JP-0078357.
 PR 18-SEP-1995; 95JP-0262062.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PI Fukuda S, Kohno K, Kunikata T, Kurimoto M, Okamura H;
 PI Taniguchi M, Tanimoto T, Torioge K, Ueshio S;
 DR WPI: 1996-252837/26.
 XX
 PT DNA encoding interferon-gamma prodn.-inducing polypeptide - useful
 PT to treat and prevent, e.g. viral disease, malignancies and immune
 PT disorders
 XX
 PS Example A-3-6; Page 37; 48pp; English.
 CC Tryptic peptides (R99561 and R99562) correspond to amino acids
 CC 79-103 and 26-43, respectively, of a novel interferon-gamma
 CC (IFN-gamma) inducer protein identified in mouse liver. The
 CC peptides were used to design PCR primers (T32405-06) that were
 CC utilised in the amplification of mouse liver cDNA, leading to
 CC the isolation of a clone (T32403) coding for mouse IFN-gamma
 CC inducer protein (R99559).
 CC
 XX
 SQ Sequence 25 AA;

Query Match 29.2%; Score 33; DB 17; Length 25;
 Best Local Similarity 35.7%; Pred. No. 72;
 Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 NWEIADQPNLEEI 14
 : : | | : : : : |
 Db 3 steemdpennid1 16

RESULT 14
 R92504
 ID R92504 standard; peptide: 25 AA.

XX AC R92504;

XX DT 30-AUG-1996 (first entry)

XX DE Interferon gamma production inducer peptide #1.

XX Interferon gamma; inducer; IFNgamma; immunocompetent cell; antiviral;
 KW antitumour; antiseptic; immunoregulatory; platelet-increasing agent;
 KW therapy; prevention; condyloma acuminatum; renal cancer; brain cancer;
 KW granuloma; mycosis fungoides; rheumatism; allergy; cytotoxicity; AIDS;
 KW killer T-cell; interleukin-2; IL-2; tumour necrosis factor; TNF;
 KW adoptive immunotherapy; monoclonal antibody.
 KW
 OS Synthetic.
 XX
 XX EP692536-A2.
 PN
 XX 17-JAN-1996.
 PD
 XX 13-JUL-1995; 95EP-0304906.
 PF
 XX 10-FEB-1995; 95JP-0045057.
 PR
 XX 14-JUL-1994; 94JP-0184162.
 PR
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA
 XX Kohno K, Kunikata T, Kurimoto M, Okamura H, Taniguchi M;
 PI
 XX Tanimoto T, Torioge K;

DR WPI: 1996-070177/08.
 XX
 PT Protein that induces gamma interferon prodn. in immuno:competent
 PT cells - used e.g. as antiviral or antitumour agent, also induces
 PT cytotoxicity of killer cells
 XX
 PS Claim 1; Page 21; 30pp; English.
 CC
 CC R92504 and R92505 represent fragments of the protein of the invention
 CC (see R92506), which induces interferon gamma (IFNgamma) production in
 CC immunocompetent cells. The protein is useful as an antiviral,
 CC antitumour, antiseptic, immunoregulatory and platelet-increasing agent.
 CC It can be used for treating AIDS, condyloma acuminatum,
 CC renal or brain cancer, granuloma, mycosis fungoides, rheumatism and
 CC allergy. The protein can also be used to induce IFNgamma production in
 CC cultured cells. The IFNgamma inducer strongly induces cytotoxicity of
 CC killer T-cells and when used with interleukin-2 (IL-2) and tumour
 CC necrosis factor (TNF), may improve the effect (or reduce side effects) of
 CC adoptive immunotherapy in tumours. The DNA encoding this protein (see
 CC T16224) can be used to produce the protein, which can then be purified
 CC (or assayed) using monoclonal antibodies.
 CC
 XX
 SQ Sequence 25 AA;

Query Match 29.2%; Score 33; DB 17; Length 25;
 Best Local Similarity 35.7%; Pred. No. 72;
 Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 NWEIADQPNLEEI 14
 : : | | : : : : |
 Db 3 steemdpennid1 16

RESULT 15
 W15702
 ID W15702 standard; peptide: 25 AA.

XX AC W15702;

XX DT 26-JAN-1998 (first entry)

XX DE Internal fragment of interferon-gamma inducer protein.

XX Interferon-gamma, IFN-gamma; antiviral; antineoplastic; radiotherapy;
 KW immunoregulatory; antitumour agent; chemotherapy; leukopenia;
 KW thrombocytopenia; immunocompetent cell; asthma; hayfever;
 KW rheumatism; interleukin; killer cell.
 KW
 OS Homo sapiens.
 XX
 XX EP767178-A1.
 PN
 XX 09-APR-1997.
 PD
 XX 26-SEP-1996; 96EP-0306997.
 PF
 XX 20-SEP-1996; 96JP-0269105.
 PR
 XX 26-SEP-1995; 95JP-0270725.
 PR
 XX 29-FEB-1996; 96JP-0067434.
 PR
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA
 XX Akita K, Fujii M, Kurimoto M, Nukeda Y, Tanimoto T;
 PI
 XX WPI: 1997-205381/19.
 DR
 XX Human protein that induces interferon-gamma prodn. in
 PT immuno:competent cells - useful for adoptive immuno:therapy of
 PT tumours and as antimicrobial agent etc.
 XX
 PS Disclosure; Page 21; 26pp; English.

CC The present sequence represents an internal fragment of a novel
 CC protein from human cells, which induces interferon-gamma (IFN gamma)
 CC production in immunocompetent cells. This protein enhances cytotoxicity
 CC of killer cells and induces their formation. It is used as an
 CC antitumor agent for antitumor immunotherapy, an antiviral (including
 CC anti-AIDS) or antibacterial agent, and in the treatment of atopic or
 CC immune system diseases, e.g. asthma, hay fever or rheumatism. When
 CC formulated with interleukin-3, it is also used to treat leukopenia and
 CC thrombocytopenia associated with radiotherapy or chemotherapy of
 CC leukemia and other cancers. When used in antitumor immunotherapy,
 CC this novel protein significantly improves the immunotherapeutic
 CC effect of interleukin-2 (IL-2), compared with use of IL-2
 CC alone, either when administered to the patient (before administration of
 CC IL-2) or by addition to the medium in which cells (intended for return to
 CC the patient) are being grown.

CC Sequence 25 AA;

Query Match 29.2%; Score 33; DB 18; Length 25;
 Best Local Similarity 35.7%; Pred. No. 72;
 Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 NMEADQPONLEI 14
 : : | | : : : :
 Db 3 steemdpennid1 16

RESULT 16

ID W24260
 AC W24260; Peptide; 25 AA.

DT 15-OCT-1997 (first entry)

DE Peptide #2 of human protein for induction of interferon-gamma.

XX Interferon gamma; immunocompetent cell; malignant tumour;

KW viral disease; bacterial infection; immune disease.

OS Homo sapiens.

XX JP09157180-A.

PN 17-JUN-1997.

PD 24-JAN-1996; 96JP-0028722.

PF 04-OCT-1995; 95JP-0279906.

PR 10-MAR-1995; 95JP-0078357.

PR 29-SEP-1995; 95JP-0274988.

XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

XX WPI; 1997-369391/34.

DR A drug containing a polypeptide which induces interferon-gamma -
 PT useful for treating e.g. malignant tumours, viral, bacterial or
 PT immune diseases

PS Disclosure; Page 10; 12pp; Japanese.

XX The sequences given in W24259-61 are peptide fragments of a protein
 CC which induces interferon-gamma production in immunocompetent cells.
 CC This protein may be used as the major component in a drug for the
 CC prevention and treatment of e.g. malignant tumours, viral diseases,
 CC bacterial infections and immune diseases. This peptide corresponds
 CC to residues 80-104 of the human protein.

XX Sequence 25 AA;

Query Match 29.2%; Score 33; DB 18; Length 25;
 Best Local Similarity 35.7%; Pred. No. 72;
 Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 NMEADQPONLEI 14
 : : | | : : : :
 Db 3 steemdpennid1 16

RESULT 17

ID Y53906
 AC Y53906; Peptide; 25 AA.

DT 13-MAR-2000 (first entry)

DE Fragment derived from a protein that induces IFN-gamma production.

XX Mouse; interferon gamma production; IFN-gamma; immunocompetent cell;

KW antiviral; immunoregulatory; antigen; mitogen;

KW IFN-gamma susceptible disease; antibacterial; antitumour;

KW blood platelet enhancing agent; hepatitis; herpes syndrome; condyloma;

KW AIDS; bacterial disease; candidiasis; malaria; solid malignant tumour;

KW renal cancer; mycosis fungoides; chronic granulomatous disease;

KW blood cell malignant tumour; adult T cell leukaemia;

KW chronic myelogenous leukaemia; malignant leukaemia; immune disease;

XX allergy; rheumatism.

OS Mus sp.

XX EP962531-A2.

PD 08-DEC-1999.

PF 10-NOV-1995; 99EP-0104104.

PR 15-NOV-1994; 94JP-0304203.

PR 23-FEB-1995; 95JP-0058240.

PR 10-MAR-1995; 95JP-0078357.

PR 18-SEP-1995; 95JP-0262062.

PR 29-SEP-1995; 95JP-0274988.

PR 10-NOV-1995; 95EP-0308055.

XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

XX Ushio S, Torigoe K, Tanimoto T, Okamura H;

XX WPI; 2000-064289/06.

DR Novel polypeptides used in the treatment of interferon-gamma

PT susceptible diseases -

PS Example A2; Page 14; 42pp; English.

XX The present sequence represents a fragment derived from a murine
 CC protein that induces interferon (IFN)-gamma production by immunocompetent
 CC cells. IFN-gamma is a protein which has antiviral, antitumor and
 CC immunoregulatory activities, and is produced by immunocompetent cells
 CC stimulated with antigens or mitogens. The present peptide was used to
 CC design probes which were used to isolate the corresponding human protein
 CC from human liver cells. The protein of the invention is used to treat
 CC IFN-gamma susceptible diseases, and also have use as a antiviral agent,
 CC antibacterial agent, antitumor agent, immunoregulatory agent and blood
 CC platelet enhancing agent. Diseases which can be treated with the
 CC protein include viral diseases such as hepatitis, herpes syndrome,
 CC condyloma, and AIDS; bacterial diseases such as Candidiasis and malaria;
 CC solid malignant tumours such as renal cancer, mycosis fungoides, and
 CC chronic granulomatous disease; blood cell malignant tumours such as
 CC adult T cell leukaemia, chronic myelogenous leukaemia, and malignant
 CC leukaemia; and immune diseases such as allergy and rheumatism.

XX Sequence 25 AA;

Query Match 29.2%; Score 33; DB 21; Length 25;
 Best Local Similarity 35.7%; Pred. No. 72;
 Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 1 NWEADOPONLEI 14
 :|||:|:|:|:
 Db 3 sfemdpennid1 16

RESULT 18

Y34179 standard; peptide: 17 AA.

Y34179;

15-NOV-1999 (first entry)

Human pre-proheparanase protein sequence fragment #6.

Human; pre-proheparanase; platelet; wound healing; angiogenesis blocker;
 inflammation; psoriasis; diabetic retinopathy; solid tumour; arthritis;
 heparin degradation; anticoagulant neutralisation; asthma; CNS disease;
 inflammatory disease; vascular restenosis; atherosclerosis; diagnosis;
 tumour growth; fibroproliferative disorder; neurodegenerative disease;
 therapy.

Homo sapiens.

WO9943830-A2.

02-SEP-1999.

18-FEB-1999; 99WO-US01489.

26-MAR-1998; 98US-0079401.

24-FEB-1998; 98US-0075706.

(PHARMA) PHARMACIA & UPJOHN CO.

Fairbanks MB, Helmricksen RL, Mildner AM;

WPI; 1999-540598/45.

New isolated platelet heparanase polypeptides, used to develop
 products for, e.g. wound healing and blocking angiogenesis

Example 3; Fig 8; 57pp; English.

This sequence represents a fragment of the human pre-proheparanase of the
 invention. The pre-proheparanase sequence was isolated from human
 platelets. The heparanase can be used for identifying agents which alter
 heparanase activity. The heparanase can be used for wound healing or for
 blocking angiogenesis or inflammation. It can be used for treating
 e.g. psoriasis, diabetic retinopathy or solid tumours, or for the
 degradation of heparin and the neutralisation of heparin's anticoagulant
 properties during surgery. Inhibitors of heparanase activity can be used
 in the treatment of arthritis, asthma, and other inflammatory diseases,
 vascular restenosis, atherosclerosis, tumour growth and progression,
 fibroproliferative disorders, and central nervous system (CNS) and
 neurodegenerative diseases. The products can also be used for detection
 and diagnosis. The purified heparanase, both recombinantly produced human
 heparanase and heparanase isolated from human platelet activity, allows
 for the convenient selection of compounds having anti-heparanase
 activity, i.e. inhibitors of heparanase activity, by measuring inhibition
 of heparanase activity. Inhibition of heparanase activity can be measured
 by blocking heparanase-mediated release of radioactive fragments from in
 vivo radiolabelled (HSPG)/heparin.

Sequence 17 AA;

Query Match 28.3%; Score 32; DB 20; Length 17;
 Best Local Similarity 40.0%; Pred. No. 67;
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 NWEADOPON 10
 :|||:|:|:|:
 Db 5 sweignepns 14

RESULT 19

Y42241 standard; peptide: 20 AA.

Y42241;

20-DEC-1999 (first entry)

Multivalent ligand Tip-18.40 peptide derivative.

Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
 endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;
 obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
 wound healing; arteriosclerosis; ischaemia; diabetes mellitus;
 blood vessel occlusion.

Synthetic.

WO9948923-A1.

30-SEP-1999.

22-MAR-1999; 99WO-US06246.

24-MAR-1998; 98US-0046985.

(CHIT-) CHILDRENS MEDICAL CENT.

(YISS) YISSUM RES & DEV CO.

Ben-Sasson SA;

WPI; 1999-601204/51.

New peptides and derived multivalent ligands based on angiogenic
 homology regions, used to inhibit or promote angiogenesis, e.g. for
 treating tumours

Claim 12; Fig 2; 76pp; English.

The present invention specifically describes peptides comprising the
 angiogenic homology region (AHR) of thrombospondin (TSP) 4, angiotatin,
 endostatin or TSP-1. Multivalent ligands based on the peptides and
 related angiogenic peptides may be anti-angiogenic, e.g. for treating
 tumours, cardiovascular disease (arteriosclerosis, ischaemia), obesity,
 osteoarthritis, duodenal ulcers, abnormal ocular vascularisation in
 diabetes, or they are proangiogenic, e.g. for promotion of wound
 healing and to stimulate neovascularisation around occluded blood
 vessels (a potential alternative to by-pass surgery or angioplasty).
 The multivalent ligands may be used in human or veterinary medicine. They
 may also be used raise peptide-specific antibodies (used for detecting
 the peptides) and to identify and isolate compounds that interact with,
 and modulate activity of, AHR. AHR are relatively small, conserved
 sequences from different angiogenic peptides that are (largely)
 responsible for biological activity. They are cheaper to prepare than
 complete proteins; may be effective at lower doses; have long-lasting
 in vivo effect and good biodistribution following oral or parenteral
 administration. The present sequence represents a multivalent ligand
 peptide derivative given in the present invention.

Sequence 20 AA;

Query Match 28.3%; Score 32; DB 20; Length 20;
 Best Local Similarity 35.7%; Pred. No. 81;

Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 WELADQPNLEEL 15
| : : : : :
Db 2 wrlshrpkdlyslv 15

RESULT 20

Y43409
Y43409 standard; peptide; 20 AA.

Y43409;

20-DEC-1999 (first entry)

Multivalent ligand Tip-18.40 peptide derivative.

Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;
obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
wound healing; arteriosclerosis; ischemic limb; ischemic myocardium;
diabetes mellitus; blood vessel occlusion.

Synthetic.

W09948924-A1.

30-SEP-1999.

23-MAR-1999; 99WO-US06334.

24-MAR-1998; 98US-0046737.

(CHIL-) CHILDRENS MEDICAL CENT.

(YISS) YISSUM RES & DEV CO.

Ben-Sasson SA;

WPI; 1999-591075/50.

New angiogenic peptide derivatives, used for treating e.g. cancer,
cardiovascular diseases, obesity, osteoarthritis, duodenal ulcers,
abnormal neovascularisation and for wound healing

Disclosure; Fig 3; 62pp; English.

The present invention specifically describes peptide derivatives
comprising an angiogenic homology region (AHR) of endostatin. The peptide
derivatives can be used for modulating angiogenesis in humans and
animals. The peptides can be used to treat a wide variety of disease
conditions, including cancer, cardiovascular diseases (e.g.
arteriosclerosis, ischemic limbs and ischemic myocardium), obesity,
osteoarthritis, duodenal ulcers, abnormal ocular neovascularisation
associated e.g. with diabetes mellitus, and to promote wound healing or
to stimulate the growth of new blood vessels to bypass, e.g. blood vessel
occlusions. The peptide derivatives can also be used for the production
of antibodies. The multivalent ligands may enable the administration of
lower doses in order to achieve therapeutic efficacy, as compared with
a univalent peptide chain. In addition, they can have long in vivo
lifetimes and good biodistribution when administered orally or
parenterally. The present sequence represents a multivalent ligand
Tip-18.40 peptide derivative given in the present invention.

Sequence 20 AA;

Query Match 28.3%; Score 32; DB 20; Length 20;

Best Local Similarity 35.7%; Pred. No. 81;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 WELADQPNLEEL 15
| : : : : :
Db 2 wrlshrpkdlyslv 15

RESULT 21

Y42236
Y42236 standard; peptide; 22 AA.

Y42236;

20-DEC-1999 (first entry)

Multivalent ligand Tip-15.40 hybrid peptide.

Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;
obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
wound healing; arteriosclerosis; ischemia; diabetes mellitus;
blood vessel occlusion.

Synthetic.

W09948923-A1.

30-SEP-1999.

22-MAR-1999; 99WO-US06246.

24-MAR-1998; 98US-0046985.

(CHIL-) CHILDRENS MEDICAL CENT.

(YISS) YISSUM RES & DEV CO.

Ben-Sasson SA;

WPI; 1999-601204/51.

New peptides and derived multivalent ligands based on angiogenic
homology regions, used to inhibit or promote angiogenesis, e.g. for
treating tumours

Claim 7; Fig 2; 76pp; English.

The present invention specifically describes peptides comprising the
angiogenic homology region (AHR) of thrombospondin (TSP) 4, angiotatin,
endostatin or TSP-1. Multivalent ligands based on the peptides and
related angiogenic peptides may be anti-angiogenic, e.g. for treating
tumours, cardiovascular disease (arteriosclerosis, ischemia), obesity,
osteoarthritis, duodenal ulcers, abnormal ocular vascularisation in
diabetes, or they are proangiogenic, e.g. for promotion of wound
healing and to stimulate neovascularisation around occluded blood
vessels (a potential alternative to by-pass surgery or angioplasty).
The multivalent ligands may be used in human or veterinary medicine. They
may also be used to raise peptide-specific antibodies (used for detecting
the peptides) and to identify and isolate compounds that interact with,
and modulate activity of, AHR. AHR are relatively small, conserved
sequences from different angiogenic peptides that are (largely)
responsible for biological activity. They are cheaper to prepare than
complete proteins; may be effective at lower doses, have long-lasting
in vivo effect and good biodistribution following oral or parenteral
administration. The present sequence represents a multivalent ligand
hybrid peptide given in the present invention.

Sequence 22 AA;

Query Match 28.3%; Score 32; DB 20; Length 22;

Best Local Similarity 35.7%; Pred. No. 90;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 WELADQPNLEEL 15
| : : : : :
Db 4 wrlshrpkdlyslv 17

Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 WELADOPONLEIL 15
| : : : : :
DB 2 WLSHPKALYSIV 15

RESULT 20

ID Y43409 standard; peptide: 20 AA.

AC Y43409;

DT 20-DEC-1999 (first entry)

DE Multivalent ligand Tip-18.40 peptide derivative.

KM Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
KM endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;
KM obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
KM wound healing; arteriosclerosis; ischaemic limb; ischaemic myocardium;
KM diabetes mellitus; blood vessel occlusion.

OS Synthetic.

PN WO9948924-A1.

PD 30-SEP-1999.

PF 23-MAR-1999; 99WO-US06334.

PR 24-MAR-1998; 98US-0046737.

PA (CHIL-) CHILDRENS MEDICAL CENT.

PA (YISS) YISSUM RES & DEV CO.

PI Ben-Sasson SA;

XX WPI; 1999-591075/50.

PT New angiogenic peptide derivatives, used for treating e.g. cancer,
PT cardiovascular diseases, obesity, osteoarthritis, duodenal ulcers,
PT abnormal neovascularisation and for wound healing

PS Disclosure: Fig 3; 62pp; English.

XX The present invention specifically describes peptide derivatives
XX comprising an angiogenic homology region (AHR) of endostatin. The peptide
XX derivatives can be used for modulating angiogenesis in humans and
XX animals. The peptides can be used to treat a wide variety of disease
XX conditions, including cancer, cardiovascular diseases (e.g.
XX arteriosclerosis, ischaemic limbs and ischaemic myocardium), obesity,
XX osteoarthritis, duodenal ulcers, abnormal ocular neovascularisation
XX associated e.g. with diabetes mellitus, and to promote wound healing or
XX to stimulate the growth of new blood vessels to bypass, e.g. blood vessel
XX occlusions. The peptide derivatives can also be used for the production
XX of antibodies. The multivalent ligands may enable the administration of
XX lower doses in order to achieve therapeutic efficacy, as compared with
XX a univalent peptide chain. In addition, they can have long in vivo
XX lifetimes and good biodistribution when administered orally or
XX parenterally. The present sequence represents a multivalent ligand
XX Tip-18.40 peptide derivative given in the present invention.

CC Sequence 20 AA;

Query Match 28.3%; Score 32; DB 20; Length 20;

Best Local Similarity 35.7%; Pred. No. 81;

Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 WELADOPONLEIL 15
| : : : : :
DB 2 WLSHPKALYSIV 15

RESULT 21

ID Y42236 standard; peptide: 22 AA.

AC Y42236;

DT 20-DEC-1999 (first entry)

DE Multivalent ligand Tip-15.40 hybrid peptide.

KM Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
KM endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;
KM obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
KM wound healing; arteriosclerosis; ischaemia; diabetes mellitus;
KM blood vessel occlusion.

OS Synthetic.

PN WO9948923-A1.

PD 30-SEP-1999.

PF 22-MAR-1999; 99WO-US06246.

PR 24-MAR-1998; 98US-0046985.

PA (CHIL-) CHILDRENS MEDICAL CENT.

PA (YISS) YISSUM RES & DEV CO.

PI Ben-Sasson SA;

XX WPI; 1999-601204/51.

PT New peptides and derived multivalent ligands based on angiogenic
PT homology regions, used to inhibit or promote angiogenesis, e.g. for
PT treating tumours

PS Claim 7; Fig 2; 76pp; English.

XX The present invention specifically describes peptides comprising the
XX angiogenic homology region (AHR) of thrombospondin (TSP) 4, angiotatin,
XX endostatin or TSP-1. Multivalent ligands based on the peptides and
XX related angiogenic peptides may be anti-angiogenic, e.g. for treating
XX tumours, cardiovascular disease (arteriosclerosis, ischaemia), obesity,
XX diabetes, or they are proangiogenic, e.g. for promotion of wound
XX healing and to stimulate neovascularisation around occluded blood
XX vessels (a potential alternative to by-pass surgery or angioplasty).
XX The multivalent ligands may be used in human or veterinary medicine. They
XX may also be used raise peptide-specific antibodies (used for detecting
XX the peptides) and to identify and isolate compounds that interact with,
XX and modulate activity of, AHR. AHR are relatively small, conserved
XX sequences from different angiogenic peptides that are (largely)
XX responsible for biological activity. They are cheaper to prepare than
XX complete proteins, may be effective at lower doses, have long-lasting
XX in vivo effect and good biodistribution following oral or parenteral
XX administration. The present sequence represents a multivalent ligand
XX hybrid peptide given in the present invention.

CC Sequence 22 AA;

Query Match 28.3%; Score 32; DB 20; Length 22;

Best Local Similarity 35.7%; Pred. No. 90;

Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 WELADOPONLEIL 15
| : : : : :
DB 4 WLSHPKALYSIV 17

RESULT 22
 Y42240 standard; peptide; 22 AA.
 AC Y42240;
 XX
 DT 20-DEC-1999 (first entry)
 DE Multivalent ligand T1p-16.40 peptide derivative.
 XX
 KW Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
 endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;
 obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
 wound healing; arteriosclerosis; ischaemia; diabetes mellitus;
 blood vessel occlusion.
 KM
 XX
 OS Synthetic.
 OS
 PN MO9948923-A1.
 XX
 PD 30-SEP-1999.
 XX
 PF 22-MAR-1999; 99WO-US06246.
 XX
 PR 24-MAR-1998; 98US-0046985.
 XX
 PA (CHIL-) CHILDRENS MEDICAL CENT.
 PA (YISS) YISSUM RES & DEV CO.
 XX
 PI Ben-Sasson SA;
 XX
 DR WPI: 1999-601204/51.
 XX
 PT New peptides and derived multivalent ligands based on angiogenic
 PT homology regions, used to inhibit or promote angiogenesis, e.g. for
 PT treating tumours -
 XX
 PS Claim 12; Fig 2; 76pp; English.
 XX
 CC The present invention specifically describes peptides comprising the
 CC angiogenic homology region (AHR) of thrombospondin (TSP) 4, angiotatin,
 CC endostatin or TSP-1. Multivalent ligands based on the peptides and
 CC related angiogenic peptides may be anti-angiogenic, e.g. for treating
 CC tumours, cardiovascular disease (arteriosclerosis, ischaemia), obesity,
 CC osteoarthritis, duodenal ulcers, abnormal ocular vascularisation in
 CC diabetes, or they are proangiogenic, e.g. for promotion of wound
 CC healing and to stimulate neovascularisation around occluded blood
 CC vessels (a potential alternative to by-pass surgery or angioplasty).
 CC The multivalent ligands may be used in human or veterinary medicine. They
 CC may also be used raise peptide-specific antibodies (used for detecting
 CC the peptides) and to identify and isolate compounds that interact with,
 CC and modulate activity of, AHR. AHR are relatively small, conserved
 CC sequences from different angiogenic peptides that are (largely)
 CC responsible for biological activity. They are cheaper to prepare than
 CC complete proteins; may be effective at lower doses; have long-lasting
 CC in vivo effect and good biodistribution following oral or parenteral
 CC administration. The present sequence represents a multivalent ligand
 CC peptide derivative given in the present invention.
 CC
 SO Sequence 22 AA;
 Query Match 28.3%; Score 32; DB 20; Length 22;
 Best Local Similarity 35.7%; Pred. No. 90;
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 OY 2 WEADQPONLEIL 15
 DB 4 WLSHNPXDLYSIV 17
 RESULT 23
 Y43404

ID Y43404 standard; peptide; 22 AA.
 AC Y43404;
 XX
 DT 20-DEC-1999 (first entry)
 DE Multivalent ligand T1p-15.40 hybrid peptide.
 XX
 KW Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
 endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;
 obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
 wound healing; arteriosclerosis; ischaemic limb; ischaemic myocardium;
 diabetes mellitus; blood vessel occlusion.
 KM
 XX
 OS Synthetic.
 OS
 PN MO9948924-A1.
 XX
 PD 30-SEP-1999.
 XX
 PF 23-MAR-1999; 99WO-US06334.
 XX
 PR 24-MAR-1998; 98US-0046737.
 XX
 PA (CHIL-) CHILDRENS MEDICAL CENT.
 PA (YISS) YISSUM RES & DEV CO.
 XX
 PI Ben-Sasson SA;
 XX
 DR WPI: 1999-591075/50.
 XX
 PT New angiogenic peptide derivatives, used for treating e.g. cancer,
 PT cardiovascular diseases, obesity, osteoarthritis, duodenal ulcers,
 PT abnormal neovascularisation and for wound healing -
 XX
 PS Disclosure; Fig 2; 62pp; English.
 XX
 CC The present invention specifically describes peptide derivatives
 CC comprising an angiogenic homology region (AHR) of endostatin. The peptide
 CC derivatives can be used for modulating angiogenesis in humans and
 CC animals. The peptides can be used to treat a wide variety of disease
 CC conditions, including cancer, cardiovascular diseases (e.g.
 CC arteriosclerosis, ischaemic limbs and ischaemic myocardium), obesity,
 CC osteoarthritis, duodenal ulcers, abnormal ocular neovascularisation
 CC associated e.g. with diabetes mellitus, and to promote wound healing or
 CC to stimulate the growth of new blood vessels to bypass, e.g. blood vessel
 CC occlusions. The peptide derivatives can also be used for the production
 CC of antibodies. The multivalent ligands may enable the administration of
 CC lower doses in order to achieve therapeutic efficacy, as compared with
 CC a univalent peptide chain. In addition, they can have long in vivo
 CC lifetimes and good biodistribution when administered orally or
 CC parenterally. The present sequence represents a multivalent ligand
 CC T1p-15.40 hybrid peptide given in the present invention.
 CC
 SO Sequence 22 AA;
 Query Match 28.3%; Score 32; DB 20; Length 22;
 Best Local Similarity 35.7%; Pred. No. 90;
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 OY 2 WEADQPONLEIL 15
 DB 4 WLSHNPXDLYSIV 17
 RESULT 24
 Y43408
 ID Y43408 standard; peptide; 22 AA.
 AC Y43408;
 XX
 DT 20-DEC-1999 (first entry)

XX Multivalent ligand Tip-16.40 peptide derivative.
 DE Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
 XX endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;
 KM obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
 KM wound healing; arteriosclerosis; ischaemic limb; ischaemic myocardium;
 KM diabetes mellitus; blood vessel occlusion.
 XX
 OS Synthetic.
 XX
 PN MO9948924-A1.
 PD 30-SEP-1999.
 XX
 PF 23-MAR-1999; 99WO-US06334.
 XX
 PR 24-MAR-1998; 98US-0046737.
 XX
 PA (CHIL-) CHILDRENS MEDICAL CENT.
 PA (YISS) YISSUM RES & DEV CO.
 PI Ben-Sasson SA;
 PI
 XX MPI; 1999-591075/50.
 DR
 XX New angiogenic peptide derivatives, used for treating e.g. cancer,
 PT cardiovascular diseases, obesity, osteoarthritis, duodenal ulcers,
 PT abnormal neovascularisation and for wound healing
 XX
 PS Disclosure; Fig 3; 62pp; English.
 XX
 CC The present invention specifically describes peptide derivatives
 CC comprising an angiogenic homology region (AHR) of endostatin. The peptide
 CC derivatives can be used for modulating angiogenesis in humans and
 CC animals. The peptides can be used to treat a wide variety of disease
 CC conditions, including cancer, cardiovascular diseases (e.g.
 CC arteriosclerosis, ischaemic limbs and ischaemic myocardium), obesity,
 CC osteoarthritis, duodenal ulcers, abnormal ocular neovascularisation
 CC associated e.g. with diabetes mellitus, and to promote wound healing or
 CC to stimulate the growth of new blood vessels to bypass, e.g. blood vessel
 CC occlusions. The peptide derivatives can also be used for the production
 CC of antibodies. The multivalent ligands may enable the administration of
 CC lower doses in order to achieve therapeutic efficacy, as compared with
 CC a univalent peptide chain. In addition, they can have long in vivo
 CC lifetimes and good biodistribution when administered orally or
 CC parenterally. The present sequence represents a multivalent ligand
 CC Tip-16.40 peptide derivative given in the present invention.
 CC
 SO Sequence 22 AA;
 XX
 Query Match 28.3%; Score 32; DB 20; Length 22;
 Best Local Similarity 35.7%; Pred. No. 90;
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 OY 2 WELADOPONLEIL 15
 | : : : : :
 DB 4 wrishrpkdlyslv 17
 XX
 RESULT 25
 Y42239
 ID Y42239 standard; peptide; 23 AA.
 XX
 AC Y42239;
 XX
 DT 20-DEC-1999 (first entry)
 XX
 DE Multivalent ligand Tip-17.40 peptide derivative.
 XX
 KM Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
 KM endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;

KM obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
 KM wound healing; arteriosclerosis; ischaemia; diabetes mellitus;
 KM blood vessel occlusion.
 XX
 OS Synthetic.
 XX
 PN MO9948923-A1.
 PD 30-SEP-1999.
 XX
 PF 22-MAR-1999; 99WO-US06246.
 XX
 PR 24-MAR-1998; 98US-0046985.
 XX
 PA (CHIL-) CHILDRENS MEDICAL CENT.
 PA (YISS) YISSUM RES & DEV CO.
 PI Ben-Sasson SA;
 PI
 XX MPI; 1999-601204/51.
 DR
 XX New peptides and derived multivalent ligands based on angiogenic
 PT homology regions, used to inhibit or promote angiogenesis, e.g. for
 PT treating tumours
 XX
 PS Claim 12; Fig 2; 76pp; English.
 XX
 CC The present invention specifically describes peptides comprising the
 CC angiogenic homology region (AHR) of thrombospondin (TSP) 4, angiotatin,
 CC endostatin or TSP-1. Multivalent ligands based on the peptides and
 CC related angiogenic peptides may be anti-angiogenic, e.g. for treating
 CC tumors, cardiovascular disease (arteriosclerosis, ischaemia), obesity,
 CC osteoarthritis, duodenal ulcers, abnormal ocular vascularisation in
 CC diabetes, or they are proangiogenic, e.g. for promotion of wound
 CC healing and to stimulate neovascularisation around occluded blood
 CC vessels (a potential alternative to by-pass surgery or angioplasty).
 CC The multivalent ligands may be used in human or veterinary medicine. They
 CC may also be used to raise peptide-specific antibodies (used for detecting
 CC the peptides) and to identify and isolate compounds that interact with,
 CC and modulate activity of, AHR. AHR are relatively small, conserved
 CC sequences from different angiogenic peptides that are (largely)
 CC responsible for biological activity. They are cheaper to prepare than
 CC complete proteins; may be effective at lower doses; have long-lasting
 CC in vivo effect and good biodistribution following oral or parenteral
 CC administration. The present sequence represents a multivalent ligand
 CC peptide derivative given in the present invention.
 CC
 SO Sequence 23 AA;
 XX
 Query Match 28.3%; Score 32; DB 20; Length 23;
 Best Local Similarity 35.7%; Pred. No. 95;
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 OY 2 WELADOPONLEIL 15
 | : : : : :
 DB 5 wrishrpkdlyslv 18

Search completed: February 5, 2001, 10:47:31
 Job time: 625 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:49:28 ; Search time 58.67 seconds
(without alignments)
23.147 Million cell updates/sec

Title: US-08-981-824-3

Perfect score: 113

Sequence: 1 NMEIADQPONLEIIMHCOR 20

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 5001

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

1: PIR66:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	27.4	17	2	B48943 phage antigenic de
2	28	24.8	18	2	F49215 urease (EC 3.5.1.5
3	27	23.9	18	2	S10452 hypothetical prote
4	27	23.9	24	2	S68391 H+-transporting AT
5	25.5	22.6	25	2	T17886 aminoglycoside pho
6	25	22.1	20	2	JT0410 bombyxin-IV chain
7	25	22.1	25	2	PN0632 biphenyl dioxygena
8	24	21.2	11	2	S21127 precorrin methyltr
9	24	21.2	19	2	PH0793 T-cell receptor al
10	24	21.2	23	2	S13268 heat shock protein
11	24	21.2	24	2	S65047 ATP-dependent clp
12	24	21.2	24	2	B23317 threonine dehydrat
13	24	21.2	25	2	S73387 alpha-2-macroglobu
14	23.5	20.8	17	2	S66198 alcohol dehydrog
15	23	20.4	9	2	A12872 transaldolase (EC
16	23	20.4	9	2	A11497 transaldolase (EC
17	23	20.4	14	2	F45037 TCR delta chain V-
18	23	20.4	15	2	PA0020 protein QAI0028 -
19	23	20.4	15	2	S51735 beta-1 receptor be
20	23	20.4	18	2	PN0149 beta-Gliadin 13 -
21	23	20.4	20	2	D42842 antifungal 2S stor
22	23	20.4	20	2	S15861 estrogen receptor
23	23	20.4	21	2	D48516 PEB1 5'-region hyp
24	23	20.4	22	2	S52357 hypothetical prote
25	23	20.4	24	2	T42258 Ser/Thr protein ph
26	23	20.4	24	2	PN0655 alkaline trypsin-1
27	23	20.4	25	2	S10850 alpha-amylase Inhl
28	23	20.4	25	2	S52124 alpha-gliadin S11
29	23	20.4	25	2	B36934 Orls 3' of mada -

30	22	19.5	6	2	B31263 dihydrofolate redu
31	22	19.5	15	2	A26228 spot 42 protein -
32	22	19.5	15	2	T46625 hypothetical prote
33	22	19.5	15	2	JN0263 antigen (clone PV1
34	22	19.5	20	2	S00493 hemocyanin chain I
35	22	19.5	20	2	S11416 ribosomal protein
36	22	19.5	21	2	I65270 collagen alpha 1(I
37	22	19.5	21	2	A60684 L-ascorbate peroxi
38	22	19.5	21	2	S61306 probable nitrate r
39	22	19.5	22	2	S55308 glutathione transf
40	22	19.5	24	2	A45336 cystic fibrosis tr
41	22	19.5	25	2	PC2204 interferon-alpha L
42	22	19.5	25	2	S22228 vitronectin - dog
43	22	19.5	25	3	P00716 glutelin 7.3/31K -
44	21.5	19.0	23	2	B29175 clostripain (EC 3.
45	21	18.6	10	2	S63696 DNA polymerase - y

ALIGNMENTS

RESULT 1
B48943
phage antigenic determinant (clone pOA79) - phage phi-197 (fragment)
C:Species: phage phi-197
C:Date: 13-Aug-1997 #sequence_revision 19-Dec-1997 #text_change 19-Dec-1997
C:Accession: B48943
R:Schouler, C.; Bouet, C.; Ritzenthaler, P.; Drouet, X.; Mata, M.
Appl. Environ. Microbiol. 58, 2479-2484, 1992
A:Title: Characterization of Lactococcus lactis phage antigens.
A:Reference number: A48943; MUID:92384563
A:Contents: Lactococcus lactis
A:Accession: B48943
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-17 <SCH>
A>Note: sequence modified after extraction from NCBI backbone
A>Note: sequence extracted from NCBI backbone (NCBIN:114551)

Query Match 27.4%; Score 31; DB 2; Length 17;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 NMEIIMH 17
DB 6 NFEELIH 13

RESULT 2
F49215
urease (EC 3.5.1.5) small chain Urea - Helicobacter mustelae (ATCC 43772) (fragment)
C:Species: Helicobacter mustelae
C:Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 30-May-1997
C:Accession: F49215
R:Rupprecht, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.
Infect. Immun. 60, 5259-5266, 1992
A:Title: Purification and characterization of the urease enzymes of Helicobacter spec
A:Reference number: A49215; MUID:93084378
A:Accession: F49215
A:Status: preliminary
A:Molecule type: Protein
A:Residues: 1-18 <TR>
A>Note: sequence extracted from NCBI backbone (NCBIP:119487)
C:Superfamily: urease 26K chain; urease 11K chain homology; urease 12K chain homology
C:Keywords: hydrolase

Query Match 24.8%; Score 28; DB 2; Length 18;
Best Local Similarity 30.0%; Pred. No. 3.6e+02;
Matches 3; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
QY 8 PONEELIMH 17

Db 3 PKELDKMLH 12

RESULT 3

hypothetical protein (bpha 5' region) - Aspergillus niger

C:Species: Aspergillus niger

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Sep-1997

C:Accession: S10452

R:van Gorm, R.F.M.; Boschloo, J.G.; Kuljshoven, A.; Lange, J.; Bos, C.J.; van Balken

submitted to the EMBL Data Library, March 1990

A:Reference number: S10452

A:Accession: S10452

A:Molecule type: DNA

A:Residues: 1-18 <VAN>

A:Cross-references: EMBL:X52521; NID:g2336; PID:g2337

Query Match

Best Local Similarity 41.7%; Score 27; DB 2; Length 18;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 4 LADQPNLEIL 15

1 MTDPPLSRQEL 12

RESULT 4

H+-transporting ATP synthase (EC 3.6.1.34) delta chain, chloroplast - Chlamydomonas rein-

N:Alternate names: ATP synthase delta chain

C:Species: Chlamydomonas reinhardtii

C>Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 17-Mar-1999

C:Accession: S68391

R:Riedler, H.R.; Schmid, R.; Leu, S.; Shavit, N.; Strotmann, H.

FEBS Lett. 377, 163-166, 1995

A:Title: Isolation of CF(0)CF(1) from Chlamydomonas reinhardtii cw15 and the N-terminal

A:Reference number: S68388; MUID:96128220

A:Accession: S68391

A:Molecule type: protein

A:Residues: 1-24 <FIE>

A:Experimental source: strain CW15

A:Genetics:

C:Superfamily: H+-transporting ATP synthase delta chain

C:Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thylak

Query Match 23.9%; Score 27; DB 2; Length 24;

Best Local Similarity 60.0%; Pred. No. 7.1e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 3 ELADQPNLEIL 12

15 ELADKRLKLE 24

RESULT 5

aminoglycoside phosphotransferase - Bacillus circulans (fragment)

C:Species: Bacillus circulans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000

C:Accession: T17886

R:Aubert-Pivert, E.; Davies, J.

Gene 147, 1-11, 1994

A:Title: Biosynthesis of butirosin in Bacillus circulans NRRL B3312: Identification by s

A:Reference number: Z18808; MUID:94374689

A:Accession: T17886

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-25 <AUB>

A:Cross-references: EMBL:L20421; NID:g304142; PID:g688437; PIDN:AAA62589.1

C:Genetics:

A:Gene: aphA4/DutA

C:Function:

A:Pathway: butirosin biosynthesis

C:Superfamily: kanamycin kinase

Query Match

Best Local Similarity 22.6%; Score 25.5; DB 2; Length 25;

Matches 6; Conservative 2; Mismatches 2; Indels 5; Gaps 1;

Db 1 NWEADQPNLEIL 15

7 NW----PEELLEIL 16

RESULT 6

bombyxin-IV chain A - silkworm

C:Species: Bombyx mori (silkworm)

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Mar-1998

C:Accession: JT0410

R:Mariyama, K.; Hietter, H.; Nagasawa, H.; Isogai, A.; Tamura, S.; Suzuki, A.; Ishiz

Agaric, Biol. Chem. 52, 3035-3041, 1988

A:Title: Isolation and primary structure of bombyxin-IV, a novel molecular species of

A:Reference number: JT0410

A:Accession: JT0410

A:Molecule type: protein

A:Residues: 1-20 <MAR>

C:Superfamily: Insulin

F:6-11/Disulfide bonds: #status predicted

F:7/Disulfide bonds: Interchain (to chain B-10) #status predicted

F:20/Disulfide bonds: Interchain (to chain B-22) #status predicted

Query Match 22.1%; Score 25; DB 2; Length 20;

Best Local Similarity 31.2%; Pred. No. 1.2e+03;

Matches 5; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Db 3 ELADQPNLEILMHC 18

5 ECCIQPCTLDVLAIVYC 20

RESULT 7

biphenyl dioxygenase (EC 1.13.-.-) A4 - Pseudomonas sp. (strain LB400) (fragment)

C:Species: Pseudomonas sp.

C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 12-Apr-1995

C:Accession: P06332

R:Hofer, B.; Ellis, L.D.; Dowling, D.N.; Timmis, K.N.

Gene 130, 47-55, 1993

A:Title: Genetic analysis of a Pseudomonas locus encoding a pathway for biphenyl/poly

A:Reference number: P06332; MUID:93345822

A:Accession: P06332

A:Molecule type: DNA

A:Residues: 1-25 <HOF>

A:Cross-references: GB:X66122

C:Genetics:

A:Gene: bphA4

C:Superfamily: toluene dioxygenase ferredoxin reductase component

C:Keywords: oxidoreductase; PCB biodegradation

Query Match 22.1%; Score 25; DB 2; Length 25;

Best Local Similarity 38.5%; Pred. No. 1.5e+03;

Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 3 ELADQPNLEIL 15

8 QLDVGSNLRDL 20

RESULT 8
S21127
precocin methyltransferase - Salmonella typhimurium
C:Species: Salmonella typhimurium
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
C:Accession: S21127
R:Recessner, C.A.; Warren, M.J.; Santander, P.J.; Atshaves, B.P.; Ozaki, S.; Stojowich, N.
FEBS Lett. 301, 73-78, 1992
A:Title: Expression of 9 Salmonella typhimurium enzymes for cobinamide synthesis. Ident
A:Reference number: S21127; MUID:93083628
A:Accession: S21127
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-11 <R0E>

Query Match 21.2% Score 24; DB 2; Length 11;
Best Local Similarity 66.7% Pred. No. 8.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 QPOMLE 12
| | | | |
Db 6 QPOTIE 11

RESULT 9
PH0793
T-cell receptor alpha chain (H2 V-alpha-7.2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PH0793
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A:Title: T cell receptor genes in a series of class I major histocompatibility complex-
allelic exclusion and antigen-specific repertoire.
A:Reference number: PH0746; MUID:92078846
A:Accession: PH0793
A:Molecule type: mRNA
A:Residues: 1-19 <CAS>
A:Cross-references: EMBL:X60898
A:Experimental source: T lymphocyte
C:Keywords: T-cell receptor

Query Match 21.2% Score 24; DB 2; Length 19;
Best Local Similarity 100.0% Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 WELA 5
| | | | |
Db 4 WELA 7

RESULT 10
S13268
heat shock protein, 90K - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 21-Nov-1998
C:Accession: S13268
R:Ichih, H.; Toyoshima, T.; Mizunuma, H.; Kobayashi, R.; Tashima, Y.
Arch. Biochem. Biophys. 282, 290-296, 1990
A:Title: Three-step purification method and characterization of the bovine brain 90-kDa
A:Reference number: S13268; MUID:91053144
A:Accession: S13268
A:Molecule type: protein
A:Residues: 1-23 <IT0>
A:Experimental source: brain
C:Superfamily: heat shock protein 90
C:Keywords: cytosol; estrogen-induced protein; heat shock; homodimer; phosphoprotein; st

Query Match 21.2% Score 24; DB 2; Length 23;
Best Local Similarity 62.5% Pred. No. 2e+03;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 6 DOPOMLEE 13
| | | | |
Db 15 DQTEYLEE 22

RESULT 11
S65047
ATP-dependent clp proteinase (EC 3.4.21.-) chain P - barley chloroplast (fragment)
C:Species: chloroplast Hordeum vulgare (barley)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999
C:Accession: S65047; S57946
R:Huebschmann, T.; Hess, W.R.; Boerner, T.
Plant Mol. Biol. 30, 109-123, 1996
A:Title: Impaired splicing of the rps12 transcript in ribosome-deficient plastids.
A:Reference number: S65047; MUID:96197402
A:Accession: S65047
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-24 <HUE>
A:Cross-references: EMBL:X89562; NID:9899304; PIDN:CA61738.1; PID:9899305
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995
C:Genetics:
A:Gene: clpP
A:Genome: chloroplast
A:Function:
A:Description: ATP-driven cleavage of proteins to small peptides
A>Note: magnesium required
C:Superfamily: ATP-dependent Clp proteinase chain P
C:Keywords: ATP; chloroplast; hydrolase; serine proteinase

Query Match 21.2% Score 24; DB 2; Length 24;
Best Local Similarity 55.6% Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 12 EELIMHCQT 20
| | | | |
Db 4 EMIDKHCDT 12

RESULT 12
B22317
threonine dehydratase (EC 4.2.1.16) - Salmonella typhimurium (fragment)
C:Species: Salmonella typhimurium
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 22-Nov-1996
C:Accession: B22317
R:Kim, S.S.; Datta, P.
Biochim. Biophys. Acta 706, 27-35, 1982
A:Title: Chemical characterization of biodegradative threonine dehydratases from two
A:Reference number: A90645; MUID:83023208
A:Accession: B22317
A:Molecule type: protein
A:Residues: 1-24 <KIM>
C:Superfamily: threonine dehydratase
C:Keywords: carbon-oxygen lyase; hydro-lyase; threonine catabolism

Query Match 21.2% Score 24; DB 2; Length 24;
Best Local Similarity 50.0% Pred. No. 2.1e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 DOPOMLEEL 15
| | | | |
Db 6 DLPVAIEDIL 15

RESULT 13
S71387
alpha-2-macroglobulin - Atlantic horseshoe crab (fragments)
C:Species: Limulus polyphemus (Atlantic horseshoe crab)
C>Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-May-1999

C:Accession: S71387
 R:DOImer, K.; Husted, L.B.; Armstrong, P.B.; Sotttrup-Jensen, L.
 FEBS Lett. 393, 37-40, 1996
 A:Title: Localisation of the major reactive lysine residue involved in the self-crosslink
 A:Reference number: S71387; MUID:96397525
 A:Accession: S71387
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-13;14-25 <DOL>
 C:Keywords: dimer; disulfide bond; proteinase inhibitor

Query Match 21.2%; Score 24; DB 2; Length 25;
 Best Local Similarity 50.0%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 NMEADOP 8
 |||
 DB 2 NMEEGVP 9

RESULT 14

S66198
 alcohol dehydrogenase (EC 1.1.1.1) class III - garden pea (fragment)
 C:Species: Pisum sativum (garden pea)
 C:Date: 14-Feb-1997 #sequence_revision 29-Aug-1997 #text_change 12-Jun-1998
 C:Accession: S66198
 R:Hjeltny, L.; Hackett, M.; Shafat, J.; Danielsson, O.; Iida, J.; Hendrickson, R.C.;
 FEBS Lett. 367, 237-240, 1995
 A:Title: Multiplicity of N-terminal structures of medium-chain alcohol dehydrogenases. M
 nzyme.
 A:Reference number: S66191; MUID:95331382
 A:Accession: S66198
 A:Molecule type: protein
 A:Residues: 1-17 <HDE>
 C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
 C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 20.8%; Score 23.5; DB 2; Length 17;
 Best Local Similarity 30.8%; Pred. No. 1.7e+03;
 Matches 4; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

OY 2 WELADOPNLEI 14
 |||
 DB 5 WE-PNKPLTIEDV 16

RESULT 15

A12872
 transaldolase (EC 2.2.1.2) I - yeast (Pichia jadinii) (fragment)
 C:Species: Pichia jadinii, Candida utilis
 C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 30-Sep-1993
 C:Accession: A12872
 R:Sun, S.C.; Joris, L.; Tsolas, O.
 Arch. Biochem. Biophys. 178, 69-78, 1977
 A:Title: Purification and crystallization of transaldolase isozyme I and evidence for di
 A:Reference number: A12872; MUID:77110646
 A:Accession: A12872
 A:Molecule type: protein
 A:Residues: 1-9 <SUN>
 C:Keywords: transferase

Query Match 20.4%; Score 23; DB 2; Length 9;
 Best Local Similarity 60.0%; Pred. No. 2e+05;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 16 MHCOT 20
 |||
 DB 3 IHCBT 7

RESULT 16
 A11497
 transaldolase (EC 2.2.1.2) III - yeast (Pichia jadinii) (fragment)

C:Species: Pichia jadinii, Candida utilis
 C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 30-Sep-1993
 C:Accession: A11497
 R:Tsolas, O.; Sun, S.C.
 Arch. Biochem. Biophys. 167, 525-533, 1975
 A:Title: Isolation of a peptide containing a histidyl-cysteinyl sequence from the a
 A:Reference number: A11497; MUID:75145197
 A:Accession: A11497
 A:Molecule type: protein
 A:Residues: 1-9 <TSO>
 C:Keywords: transferase

Query Match 20.4%; Score 23; DB 2; Length 9;
 Best Local Similarity 60.0%; Pred. No. 2e+05;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 16 MHCOT 20
 |||
 DB 3 IHCBT 7

RESULT 17

F49037
 TOR delta chain V-D-J region - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
 C:Accession: F49037
 R:Bezguerra, A.; Wilde, D.B.; McConnell, T.J.; Sturmhofel, K.; Valas, R.B.; Shevach, E
 Eur. J. Immunol. 22, 491-498, 1992
 A:Title: Mouse autoreactive gamma/delta T cells. II. Molecular characterization of th
 A:Reference number: F49037; MUID:92164730
 A:Accession: F49037
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-14 <EQ>
 A:Cross-references: GB:S90654; NID:g246298; PIDN:ABR21552.1; PID:g246299
 A:Experimental source: dendritic epidermal T-cell lines
 A:Note: sequence extracted from NCBI backbone (NCBI:90654, NCBI:90666)

Query Match 20.4%; Score 23; DB 2; Length 14;
 Best Local Similarity 60.0%; Pred. No. 1.7e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 WELAD 6
 |||
 DB 4 WELSE 8

RESULT 18

PA0020
 protein OAI00028 - Arabidopsis thaliana (fragment)
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 06-Jun-1997
 C:Accession: PA0020
 R:Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
 submitted to JRPID, July 1994
 A:Description: Separation and characterization of Arabidopsis proteins by two-dimensi
 A:Reference number: PA0001
 A:Accession: PA0020
 A:Molecule type: protein
 A:Residues: 1-15 <KAM>
 A:Experimental source: callus

Query Match 20.4%; Score 23; DB 2; Length 15;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 QPONEE 13
 Db 4 QPARMEE 10

RESULT 19

S51735
 T-cell receptor beta-chain joining region - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 05-Nov-1999
 C:Accession: S51735
 R:Duhovic-Bello, I.; Steinhilber, A.; Ziegler, A.G.; Schendel, D.J.
 submitted to the EMBL Data Library, November 1993
 A:Reference number: S51732
 A:Accession: S51735
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-15 <DUP>
 A:Cross-references: EMBL:Z28344; NID:9607122; PIDN:CAA82198.1; PID:9607123
 C:Keywords: T-cell receptor

Query Match 20.4% Score 23; DB 2; Length 15;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 LADOPON 10
 Db 8 LGNOPH 14

RESULT 20

PN0149
 beta-Gliadin 13 - Aegilops longissima (fragment)
 C:Species: Aegilops longissima
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C:Accession: PN0149
 R:Odintsova, T.I.; Egorov, T.A.
 Biokhimiya 55, 509-516, 1990
 A:Title: N-Terminal sequences of omega-gliadins of Aegilops longissima: On the origin of
 A:Reference number: PN0146; MUID:90283493
 A:Accession: PN0149
 A:Molecule type: protein
 A:Residues: 1-18 <ODI>
 A:Experimental source: strain K-202
 C:Superfamily: gliadin

Query Match 20.4% Score 23; DB 2; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 OPON 10
 Db 9 QPON 12

RESULT 21

D42842
 antifungal 2S storage albumin large chain - radish (fragment)
 C:Species: Raphanus sativus (radish)
 C:Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
 C:Accession: D42842
 R:Terras, F.R.; Schoofs, H.M.; De Bolle, M.F.; Van Leuven, F.; Rees, S.B.; Vanderleyden, J.
 Biol. Chem. 267, 15301-15309, 1992
 A:Title: Analysis of two novel classes of plant antifungal proteins from radish (Raphanus)
 A:Reference number: A42842; MUID:92346373
 A:Accession: D42842
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-20 <TER>
 A:Experimental source: seed
 A:Note: sequence extracted from NCBI backbone (NCBI:109925)

C:Superfamily: wheat alpha-amylase inhibitor

Query Match 20.4% Score 23; DB 2; Length 20;
 Best Local Similarity 36.4%; Pred. No. 2.5e+03;
 Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 8 PONEILMHC 18
 Db 4 PQRPPLQOC 14

RESULT 22

S15861
 estrogen receptor - pig (fragments)
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 30-May-1997
 C:Accession: S15861
 R:Thiele, H.H.; Jungblut, P.W.; Jakob, F.
 Biochem. J. 276, 709-714, 1991
 A:Title: The proton-driven dissociation of oestradiol-receptor dimers as a preparative
 A:Reference number: S15861; MUID:91291128
 A:Accession: S15861
 A:Molecule type: protein
 A:Residues: 1-20 <BIO>
 C:Keywords: steroid hormone receptor

Query Match 20.4% Score 23; DB 2; Length 20;
 Best Local Similarity 40.0%; Pred. No. 2.5e+03;
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 NWEIADOPON 10
 Db 7 NWAKLHAPT 16

RESULT 23

D48518
 PEB1 5'-region hypothetical protein A - Campylobacter jejuni (fragment)
 C:Species: Campylobacter jejuni
 C:Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 19-May-1995
 C:Accession: D48518
 R:Pei, Z.; Blaser, M.J.
 J. Biol. Chem. 268, 18717-18725, 1993
 A:Title: PEB1, the major cell-binding factor of Campylobacter jejuni, is a homolog of
 A:Reference number: A48518; MUID:93366784
 A:Accession: D48518
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-21 <PEI>
 A:Cross-references: GB:L13662

Query Match 20.4% Score 23; DB 2; Length 21;
 Best Local Similarity 50.0%; Pred. No. 2.6e+03;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 7 OPONEEI 14
 Db 3 KPMILKEI 10

RESULT 24

S52357
 hypothetical protein - human
 C:Species: Homo sapiens (man)
 C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999
 C:Accession: S52357
 R:Agriokastilis, A.; Leversha, M.A.; Ferguson-Smith, M.; Moschonas, M.K.
 submitted to the EMBL Data Library, March 1993
 A:Description: A cosmid clone mapped to human chromosome 11p15 detects a Tag I restr

A:Reference number: S52355
A:Accession: S52357
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <AGR>
A:Cross-references: EMBL:X72881; NID:g667002; PIDN:CAA51393.1; PID:g667004

Query Match
Best Local Similarity 20.4%; Score 23; DB 2; Length 22;
Matches 5; Conservativity 50.0%; Pred. No. 2.7e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 4 LADQPNLEE 13
| : | | | |
Db 12 LGEORANTEE 21

RESULT 25

T42258
Ser/Thr protein phosphatase homolog - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: T42258
R:Zeke, T.; Gergely, P.; Dombradi, V.
submitted to the EMBL Data Library, July 1996
A:Description: The catalytic subunits of Ser/Thr protein phosphatases from Caenorhabditis
A:Reference number: 222131
A:Accession: T42258
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-24 <ZEK>
A:Cross-references: EMBL:277734; PIDN:CAB01293.1
C:Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprotein

Query Match
Best Local Similarity 20.4%; Score 23; DB 2; Length 24;
Matches 6; Conservativity 50.0%; Pred. No. 3e+03;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 7 QPONEEILMHC 18
| : | | | |
Db 1 QSIHLRLIMTC 12

Search completed: February 5, 2001, 10:49:29
Job time: 742 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:55:58 ; Search time 32.57 Seconds
(without alignments)
19.831 Million cell updates/sec

Title: US-08-981-824-3

Perfect score: 113
Sequence: 1 NMEIADQPONLEILMHCQT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 1436

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	22.1	23	1	UDP_LACCA
2	24	21.2	24	1	CLPP_HORVU
3	24	21.2	24	1	THD2_SALT
4	23	20.4	9	1	TAL1_PICJA
5	23	20.4	9	1	TAL3_PICJA
6	23	20.4	20	1	OXLA_AGRH
7	22	19.5	19	1	DHAB_COMTE
8	22	18.6	10	1	GN2_CHEPR
9	21	18.6	10	1	MOSQ_CICJA
10	21	18.6	16	1	MMPX_SOLTU
11	21	18.6	21	1	RL5_HALME
12	21	18.6	22	1	RL5_HALVO
13	21	18.6	23	1	IAPP_LIEPU
14	21	18.6	23	1	RL5_HALHA
15	20	17.7	15	1	FKB7_PINPS
16	20	17.7	15	1	GLN2_PINPS
17	20	17.7	15	1	VORA_MERTM
18	20	17.7	16	1	LEO5_BIOGL
19	20	17.7	17	1	EEG_THEO
20	20	17.7	19	1	DCAM_ACACA
21	20	17.7	19	1	ETFA_CICJA
22	20	17.7	19	1	OXLA_OPHHA
23	20	17.7	20	1	DETS_RAT
24	20	17.7	20	1	PSAL_SYNVU
25	20	17.7	24	1	RAN_XENLA
26	19.5	17.3	18	1	NPA_BOVIN
27	19	16.8	10	1	GNL1_ALIMI
28	19	16.8	15	1	RBS_PHYRA
29	19	16.8	17	1	BOL2_MESPE
30	19	16.8	17	1	BOL4_MESPE
31	19	16.8	17	1	GSX_PINPS
32	19	16.8	20	1	COX2_ONCMY
33	19	16.8	20	1	CPXX_RHCRH

34	19	16.8	21	1	YFLA_METVO	P42017 methanococ
35	19	16.8	24	1	IRBP_SHEEP	P12653 ovis aries
36	19	16.8	25	1	UBLI_BOVIN	P23356 bos taurus
37	18	15.9	14	1	KARA_BROPL	P22432 bromelia pl
38	18	15.9	14	1	MAST_PAPID	P42716 parapolybia
39	18	15.9	14	1	MAST_VESXA	P01515 vespa xanth
40	18	15.9	15	1	ITRB_ALBUJ	P29277 albizzia ju
41	18	15.9	16	1	CA19_PRUSE	P29259 prunus sero
42	18	15.9	16	1	CAT9_FASHE	P80533 fasciola he
43	18	15.9	17	1	BOL3_MESPE	P07494 megabombus
44	18	15.9	17	1	TPIS_PINPS	P81665 pinus pinas
45	18	15.9	18	1	MCRB_METTE	P22949 methanosarc

ALIGNMENTS

RESULT	ID	Query Match	Length	ID	Description
1	UDP_LACCA	22.1%	23	1	UDP_LACCA
2	CLPP_HORVU	21.2%	24	1	CLPP_HORVU
3	THD2_SALT	21.2%	24	1	THD2_SALT
4	TAL1_PICJA	20.4%	9	1	TAL1_PICJA
5	TAL3_PICJA	20.4%	9	1	TAL3_PICJA
6	OXLA_AGRH	20.4%	20	1	OXLA_AGRH
7	DHAB_COMTE	19.5%	19	1	DHAB_COMTE
8	GN2_CHEPR	18.6%	10	1	GN2_CHEPR
9	MOSQ_CICJA	18.6%	10	1	MOSQ_CICJA
10	MMPX_SOLTU	18.6%	16	1	MMPX_SOLTU
11	RL5_HALME	18.6%	21	1	RL5_HALME
12	RL5_HALVO	18.6%	22	1	RL5_HALVO
13	IAPP_LIEPU	18.6%	23	1	IAPP_LIEPU
14	RL5_HALHA	18.6%	23	1	RL5_HALHA
15	FKB7_PINPS	17.7%	15	1	FKB7_PINPS
16	GLN2_PINPS	17.7%	15	1	GLN2_PINPS
17	VORA_MERTM	17.7%	15	1	VORA_MERTM
18	LEO5_BIOGL	17.7%	16	1	LEO5_BIOGL
19	EEG_THEO	17.7%	17	1	EEG_THEO
20	DCAM_ACACA	17.7%	19	1	DCAM_ACACA
21	ETFA_CICJA	17.7%	19	1	ETFA_CICJA
22	OXLA_OPHHA	17.7%	19	1	OXLA_OPHHA
23	DETS_RAT	17.7%	20	1	DETS_RAT
24	PSAL_SYNVU	17.7%	20	1	PSAL_SYNVU
25	RAN_XENLA	17.7%	24	1	RAN_XENLA
26	NPA_BOVIN	17.3%	18	1	NPA_BOVIN
27	GNL1_ALIMI	16.8%	10	1	GNL1_ALIMI
28	RBS_PHYRA	16.8%	15	1	RBS_PHYRA
29	BOL2_MESPE	16.8%	17	1	BOL2_MESPE
30	BOL4_MESPE	16.8%	17	1	BOL4_MESPE
31	GSX_PINPS	16.8%	17	1	GSX_PINPS
32	COX2_ONCMY	16.8%	20	1	COX2_ONCMY
33	CPXX_RHCRH	16.8%	20	1	CPXX_RHCRH

RESULT	ID	Query Match	Length	ID	Description
1	UDP_LACCA	22.1%	23	1	UDP_LACCA
2	CLPP_HORVU	21.2%	24	1	CLPP_HORVU
3	THD2_SALT	21.2%	24	1	THD2_SALT
4	TAL1_PICJA	20.4%	9	1	TAL1_PICJA
5	TAL3_PICJA	20.4%	9	1	TAL3_PICJA
6	OXLA_AGRH	20.4%	20	1	OXLA_AGRH
7	DHAB_COMTE	19.5%	19	1	DHAB_COMTE
8	GN2_CHEPR	18.6%	10	1	GN2_CHEPR
9	MOSQ_CICJA	18.6%	10	1	MOSQ_CICJA
10	MMPX_SOLTU	18.6%	16	1	MMPX_SOLTU
11	RL5_HALME	18.6%	21	1	RL5_HALME
12	RL5_HALVO	18.6%	22	1	RL5_HALVO
13	IAPP_LIEPU	18.6%	23	1	IAPP_LIEPU
14	RL5_HALHA	18.6%	23	1	RL5_HALHA
15	FKB7_PINPS	17.7%	15	1	FKB7_PINPS
16	GLN2_PINPS	17.7%	15	1	GLN2_PINPS
17	VORA_MERTM	17.7%	15	1	VORA_MERTM
18	LEO5_BIOGL	17.7%	16	1	LEO5_BIOGL
19	EEG_THEO	17.7%	17	1	EEG_THEO
20	DCAM_ACACA	17.7%	19	1	DCAM_ACACA
21	ETFA_CICJA	17.7%	19	1	ETFA_CICJA
22	OXLA_OPHHA	17.7%	19	1	OXLA_OPHHA
23	DETS_RAT	17.7%	20	1	DETS_RAT
24	PSAL_SYNVU	17.7%	20	1	PSAL_SYNVU
25	RAN_XENLA	17.7%	24	1	RAN_XENLA
26	NPA_BOVIN	17.3%	18	1	NPA_BOVIN
27	GNL1_ALIMI	16.8%	10	1	GNL1_ALIMI
28	RBS_PHYRA	16.8%	15	1	RBS_PHYRA
29	BOL2_MESPE	16.8%	17	1	BOL2_MESPE
30	BOL4_MESPE	16.8%	17	1	BOL4_MESPE
31	GSX_PINPS	16.8%	17	1	GSX_PINPS
32	COX2_ONCMY	16.8%	20	1	COX2_ONCMY
33	CPXX_RHCRH	16.8%	20	1	CPXX_RHCRH

RESULT	ID	Query Match	Length	ID	Description
1	UDP_LACCA	22.1%	23	1	UDP_LACCA
2	CLPP_HORVU	21.2%	24	1	CLPP_HORVU
3	THD2_SALT	21.2%	24	1	THD2_SALT
4	TAL1_PICJA	20.4%	9	1	TAL1_PICJA
5	TAL3_PICJA	20.4%	9	1	TAL3_PICJA
6	OXLA_AGRH	20.4%	20	1	OXLA_AGRH
7	DHAB_COMTE	19.5%	19	1	DHAB_COMTE
8	GN2_CHEPR	18.6%	10	1	GN2_CHEPR
9	MOSQ_CICJA	18.6%	10	1	MOSQ_CICJA
10	MMPX_SOLTU	18.6%	16	1	MMPX_SOLTU
11	RL5_HALME	18.6%	21	1	RL5_HALME
12	RL5_HALVO	18.6%	22	1	RL5_HALVO
13	IAPP_LIEPU	18.6%	23	1	IAPP_LIEPU
14	RL5_HALHA	18.6%	23	1	RL5_HALHA
15	FKB7_PINPS	17.7%	15	1	FKB7_PINPS
16	GLN2_PINPS	17.7%	15	1	GLN2_PINPS
17	VORA_MERTM	17.7%	15	1	VORA_MERTM
18	LEO5_BIOGL	17.7%	16	1	LEO5_BIOGL
19	EEG_THEO	17.7%	17	1	EEG_THEO
20	DCAM_ACACA	17.7%	19	1	DCAM_ACACA
21	ETFA_CICJA	17.7%	19	1	ETFA_CICJA
22	OXLA_OPHHA	17.7%	19	1	OXLA_OPHHA
23	DETS_RAT	17.7%	20	1	DETS_RAT
24	PSAL_SYNVU	17.7%	20	1	PSAL_SYNVU
25	RAN_XENLA	17.7%	24	1	RAN_XENLA
26	NPA_BOVIN	17.3%	18	1	NPA_BOVIN
27	GNL1_ALIMI	16.8%	10	1	GNL1_ALIMI
28	RBS_PHYRA	16.8%	15	1	RBS_PHYRA
29	BOL2_MESPE	16.8%	17	1	BOL2_MESPE
30	BOL4_MESPE	16.8%	17	1	BOL4_MESPE
31	GSX_PINPS	16.8%	17	1	GSX_PINPS
32	COX2_ONCMY	16.8%	20	1	COX2_ONCMY
33	CPXX_RHCRH	16.8%	20	1	CPXX_RHCRH

```

DE (FRAGMENT).
GN CLPP.
OS Hordeum vulgare (Barley).
OG Chloroplast.
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
CC Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-CV. HAISA;
RX MEDLINE-96197402; PubMed=8616228;
RA Hubschmann T.; Hess W.; Boerner T.;
RT "Impaired splicing of the rps12 transcript in ribosome-deficient
  plants."
RL Plant Mol. Biol. 30:109-123(1996).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS TO SMALL PEPTIDES IN
CC THE PRESENCE OF ATP AND MAGNESIUM. ALPHA-CASEIN IS THE USUAL TEST
CC SUBSTRATE. IN THE ABSENCE OF ATP, ONLY OLIGOPEPTIDES SHORTER THAN
CC FIVE RESIDUES ARE CLEAVED (SUCH AS SUCCINY-L-LEU-TYR-L-NHMEC, AND
CC L-LEU-TYR-L-LEU-TYR-TYR, IN WHICH THE CLEAVAGE OF THE -TYR-L-LEU-
CC AND -TYR-L-TYR- BOND ALSO OCCURS).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S14; ALSO KNOWN AS CLPP
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: X89562; CA61738.1;
CC INTERPRO: IPR001907;
CC PROSITE: PS00381; CLP_PROTEASE_SER; PARTIAL.
CC PROSITE: PS00382; CLP_PROTEASE_HIS; PARTIAL.
CC HYDROLASE; Serine protease; Chloroplast.
CC NON_TER 1
CC SEQUENCE 24 AA; 2941 MW; 715EB01FE3E073CE CRC64;
SQ

```

Query Match 21.2%; Score 24; DB 1; Length 24;
 Best Local Similarity 55.6%; Pred. No. 8.5e+02;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 12 EELMHQOT 20
DB 4 EMDKHCDF 12

```

RESULT 3

```

THD2_SALTY STANDARD; PRT; 24 AA.
AC P11954;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1993 (Rel. 31, Last annotation update)
DE THREONINE DEHYDRATASE CATABOLIC (EC 4.2.1.16) (THREONINE DEAMINASE)
DE (FRAGMENT).
GN TDCB.
OS Salmonella typhimurium.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Salmonella.
RN [1]
RN SEQUENCE.
RX MEDLINE-83023208; PubMed-6751404;
RA Kim S.S.; Datta P.;
RT "Chemical characterization of biodegradative threonine dehydratases
  from two enteric bacteria."
RT Biochim. Biophys. Acta 706:27-35(1982).
CC -1- CATALYTIC ACTIVITY: L-THREONINE + H(2)O -> 2-OXOBUTANOATE + NH(3)
CC + H(2)O.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.

```

```

CC -1- ENZYME REGULATION: EACH PROTEIN MOLECULE CAN BIND UP TO FOUR
CC MOLECULES OF AMP, WHICH ACTS AS AN ALLOSTERIC ACTIVATOR TO THE
CC ENZYME. THE ENZYME IS ALSO INHIBITED BY ALPHA-KETO ACIDS AND OTHER
CC CATABOLITES. IT IS ALLOSTERICALLY INHIBITED BY ISOLEUCINE
CC AND ALLOSTERICALLY ACTIVATED BY VALINE.
CC -1- PATHWAY: THREONINE CATABOLISM.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: TO THREONINE DEHYDRATASE BIOSYNTHETIC AND SERINE
CC DEHYDRATASE.
CC PIR: B22317; B22317.
CC STGENE: SG10390; TDCB.
DR INTERPRO: IPR000634;
DR PROSITE: PS00165; DEHYDRATASE_SER_THR; PARTIAL.
KW Lyase; Pyridoxal phosphate; Allosteric enzyme.
FT NON_TER 24
FT SEQUENCE 24 AA; 2710 MW; 57BE403BF2C72AF3 CRC64;
SQ

```

Query Match 21.2%; Score 24; DB 1; Length 24;
 Best Local Similarity 50.0%; Pred. No. 8.5e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

OY 6 DOPQLEEL 15
DB 6 DLPVAIEDIL 15

```

RESULT 4

```

TAL3_PICJA STANDARD; PRT; 9 AA.
AC P1740;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE TRANSALDOLASE I (EC 2.2.1.2) (FRAGMENT).
OS Pichia jadinii (Yeast) (Candida utilis).
CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
CC Saccharomycetaceae; Pichia.
RN [1]
RN SEQUENCE.
RX MEDLINE-77110646; PubMed-556924;
RA Sun S.C.; Joris L.; Teolas O.;
RT "Purification of crystallization of transaldolase isozyme I and
  evidence for different genetic origin of isozymes I and III in
  Candida utilis."
RL Arch. Biochem. Biophys. 178:69-78(1977).
CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF
CC METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.
CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
CC 3-PHOSPHATE -> D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.
CC -1- PATHWAY: NONOXIDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.
CC -1- INTERPRO: IPR001585;
CC PIR: A12872; A12872.
DR INTERPRO: PS00958; TRANSALDOLASE_2; PARTIAL.
DR PROSITE: PS00954; TRANSALDOLASE_1; PARTIAL.
KW Transferase; Pentose shunt.
FT NON_TER 1
FT NON_TER 1
FT SEQUENCE 9 AA; 1008 MW; 274E31AFOEB1E058 CRC64;
SQ

```

Query Match 20.4%; Score 23; DB 1; Length 9;
 Best Local Similarity 60.0%; Pred. No. 8.9e+04;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

OY 16 MHCOT 20
DB 3 IHCBT 7

```

RESULT 5

```

TAL3_PICJA

```

ID TAL2_PICJA STANDARD: PRT: 9 AA.
 AC P1741:
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE TRANSALDOLASE III (EC 2.2.1.2) (FRAGMENT).
 OS Pichia jadinii (Yeast) (Candida utilis).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Pichia.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-75145197; PubMed-1092268;
 RA Tsoias O., Sun S.C.;
 RT "Isolation of a peptide containing a histidinyl-cysteinyl sequence
 from the active center of transaldolase."
 RL Arch. Biochem. Biophys. 167:525-533(1975).
 CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF
 METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.
 CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
 3-PHOSPHATE -> D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.
 CC -1- PATHWAY: NONOXIDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.
 DR INTERPRO: IPR001565;
 DR PROSITE: PS00558; TRANSALDOLASE_2; PARTIAL.
 DR PROSITE: PS01054; TRANSALDOLASE_1; PARTIAL.
 KW TRANSFERASE: Pentose shunt.
 FT NON-TER 1 1
 FT SEQUENCE 9 AA; 1033 MW; 325A31A4EB1E058 CRC64;

Query Match 20.4%; Score 23; DB 1; Length 9;
 Best Local Similarity 60.0%; Pred. No. 8.9e+04;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 16 MHQT 20
 DB 3 INCHT 7

RESULT 6
 OXLA_AGRH STANDARD: PRT: 20 AA.
 AC P81382;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE L-AMINO ACID OXIDASE (EC 1.4.3.2) (LAO) (LAO) (FRAGMENT).
 OS Agkistrodon rhodostoma (Malayan pit viper) (Calloselasma rhodostoma).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Viperidae; Crotalinae; Calloselasma.
 RN [1]
 RP SEQUENCE.
 RX TISSUE-VEINOM;
 RX MEDLINE-94361525; PubMed-8060286;
 RA Ponnudurai G., Chung M.C.M., Tan N.-H.;
 RT Purification and properties of the L-amino acid oxidase from Malayan
 pit viper (Calloselasma rhodostoma) venom."
 RL Arch. Biochem. Biophys. 313:373-378(1994).
 CC -1- FUNCTION: HAS CYTOTOXIC ACTIVITY. THIS PROTEIN HAS A PH OPTIMUM
 OF 9.0. A DETERMINED PI OF 4.4 AND IS TEMPERATURE STABLE.
 CC -1- CATALYTIC ACTIVITY: A L-AMINO ACID + H(2)O + O(2) -> A 2-OXO ACID +
 NH(3) + H(2)O(2).
 CC -1- COFACTOR: FAD.
 CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
 CC STRONG, TO MOUSE FIG-1.
 KW Oxidoreductase; Flavoprotein; FAD; Venom.
 FT NON-TER 20 20
 FT SEQUENCE 20 AA; 2443 MW; 96260598D7F7D07 CRC64;

Query Match 20.4%; Score 23; DB 1; Length 20;
 Best Local Similarity 55.6%; Pred. No. 1e+03;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 OPONLEEL 15
 DB 12 QENNEEEL 20

RESULT 7
 DHAB_COMTE STANDARD: PRT: 19 AA.
 AC P80704;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE ALDEHYDE DEHYDROGENASE, BETA CHAIN (EC 1.2.99.3) (ALDH) (FRAGMENT).
 OS Comamonas testosteroni (Pseudomonas testosteroni).
 OC Bacteria; Proteobacteria; Beta subdivision; Comamonadaceae; Comamonas.
 RN [1]
 RP SEQUENCE.
 RC STRAIN-ATCC 15667;
 RA Luykx D.M.A.M., Kim S.W., de Vries S., Duine J.A.;
 RL Submitted (JUL-1996) to the SWISS-PROT data bank.
 CC -1- CATALYTIC ACTIVITY: AN ALDEHYDE + ACCEPTOR + H(2)O -> AN ACID +
 REDUCED ACCEPTOR.
 CC -1- COFACTOR: MOLYBDENUM.
 CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA AND A GAMMA
 CHAIN.
 KW Oxidoreductase; Molybdenum.
 FT NON-TER 19 19
 FT SEQUENCE 19 AA; 2096 MW; 89BD67DAD05A212E CRC64;

Query Match 19.5%; Score 22; DB 1; Length 19;
 Best Local Similarity 42.9%; Pred. No. 1.4e+03;
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 PONTLEET 14
 DB 9 PRTIDEV 15

RESULT 8
 GON2_CHEPR STANDARD: PRT: 10 AA.
 AC P80678;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GONADOLIBERIN II (GONADOTROPIN-RELEASING HORMONE II) (GNRH-II)
 DE (LUTIBERIN II).
 OS Cheliosoma productum.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
 OC Corellidae; Chelyosoma.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-96413669; PubMed-8816823;
 RA Powell J.F.F., Reeska-Skinner S.M., Prakash M.O., Fischer W.H.,
 RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
 RT "Two new forms of gonadotropin-releasing hormone in a protochordate
 and the evolutionary implications."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 FOLLICLE-STIMULATING HORMONES.
 CC -1- SUBUNIT: HOMODIMER, LINKED BY DISULFIDE BRIDGE.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
 THE GONADOTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
 THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
 CC -1- MASS SPECTROMETRY: MW-1117.52; METHOD-MALDI.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR INTERPRO: IPR002012;

DR PROSITE; PS00473; GNRH; 1.
 KW Hormone; Amidation.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 6 6 INTERCHAIN.
 FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
 SQ SEQUENCE 10 AA: 1135 MW: 284538DLEEB735A3 CRC64;

Query Match 18.6%; Score 21; DB 1; Length 10;
 Best Local Similarity 37.5%; Pred. No. 9.6e+02;
 Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 NWEIADOP 8
 : : : :
 Db 2 HWSLCHAP 9

RESULT 9
 MOSQ_CLYXA STANDARD; PRT; 10 AA.
 ID MOSQ_CLYXA
 AC P19962;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE [GLN-6]-MOSACT.
 OS Clypeaster japonicus (sand dollar).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidae; Euechinoidea; Gnathostomata; Clypeasteroidea;
 CN Clypeasteridae; Clypeaster.
 [1]
 RP SEQUENCE.
 RC TISSUE-EGG JELLY.
 RA Suzuki N., Kurita M., Yoshino K., Kajitara H., Nomura K., Yamaguchi M.;
 RT "Purification and structure of mosact and its derivatives from the
 egg jelly of the sea urchin Clypeaster japonicus";
 RL Zool. Sci. 4:649-656(1987).
 CC -1- FUNCTION: STIMULATES SPERM RESPIRATION AND MOTILITY.
 DR PIR: JN0025; JN0025. 1019 MW: 9AFB032456DC5BA CRC64;
 SQ SEQUENCE 10 AA: 1019 MW: 9AFB032456DC5BA CRC64;

Query Match 18.6%; Score 21; DB 1; Length 10;
 Best Local Similarity 57.1%; Pred. No. 9.6e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 ADQOPNL 11
 : : : : :
 Db 2 SDSAONL 8

RESULT 10
 MPX_SOLTU STANDARD; PRT; 16 AA.
 ID MPX_SOLTU
 AC P80501;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE UNIDENTIFIED MITOCHONDRIAL MATRIX PROTEIN (FRAGMENT).
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Solanum.
 [1]
 RP SEQUENCE.
 RC TISSUE-TUBER;
 RX MEDLINE-97077345; PubMed-8919912;
 RA Jansch L., Kruff V., Schmitz U.K., Braun H.P.;
 RT "New insights into the composition, molecular mass and stoichiometry
 of the protein complexes of plant mitochondria";
 RL Plant J. 9:357-368(1996).
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
 KW Mitochondrion.
 FT NON_TER 16 16

SQ SEQUENCE 16 AA: 1768 MW: C58D4DB48A18B8D CRC64;

Query Match 18.6%; Score 21; DB 1; Length 16;
 Best Local Similarity 50.0%; Pred. No. 1.6e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 ELADQPN 10
 : : : : :
 Db 7 ELVEKGN 14

RESULT 11
 RL5_HALME STANDARD; PRT; 21 AA.
 ID RL5_HALME
 AC P50557;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L5P (HME15) (FRAGMENT).
 GN RPL5P.
 OS Halobacterium mediterranei (Halobacterax mediterranei).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Halobacter.
 [1]
 RP SEQUENCE.
 RC STRAIN-DSM 1411;
 RX MEDLINE-94229075; PubMed-8174557;
 RA McDougall J., Wittmann-Liebold B.;
 RT "Comparative analysis of the protein components from 5S rRNA-protein
 complexes of halophilic archaeobacteria";
 RL Eur. J. Biochem. 221:779-785(1994).
 CC -1- SIMILARITY: BELONGS TO THE L5P FAMILY OF RIBOSOMAL PROTEINS.
 DR INTERPRO: IPR002132;
 DR PIR: G33084; G33084.
 DR PIR: PFO0281; RIBOSOMAL_L5; PARTIAL.
 DR PROSITE; PS00358; RIBOSOMAL_L5; PARTIAL.
 KW Ribosomal protein.
 FT NON_TER 21
 SQ SEQUENCE 21 AA: 2497 MW: 56EB7371B2A13F71 CRC64;

Query Match 18.6%; Score 21; DB 1; Length 21;
 Best Local Similarity 28.6%; Pred. No. 2.2e+03;
 Matches 2; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 11 LEETLMH 17
 : : : : :
 Db 13 IEKVYVH 19

RESULT 12
 RL5_HALVO STANDARD; PRT; 22 AA.
 ID RL5_HALVO
 AC P50559;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L5P (HVOL5) (FRAGMENT).
 GN RPL5P.
 OS Halobacterium volcanii (Halobacterax volcanii).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Halobacter.
 [1]
 RP SEQUENCE.
 RC STRAIN-DSM 3757;
 RX MEDLINE-94229075; PubMed-8174557;
 RA McDougall J., Wittmann-Liebold B.;
 RT "Comparative analysis of the protein components from 5S rRNA-protein
 complexes of halophilic archaeobacteria";
 RL Eur. J. Biochem. 221:779-785(1994).
 CC -1- SIMILARITY: BELONGS TO THE L5P FAMILY OF RIBOSOMAL PROTEINS.
 DR INTERPRO: IPR002132;
 DR PIR: H33084; H33084.
 DR PIR: PFO0281; RIBOSOMAL_L5; 1.

DR PROSITE: PS00358; RIBOSOMAL_L5; PARTIAL.
 KW RIBOSOMAL protein.
 FT NON_TER 22
 SQ SEQUENCE 22 AA; 2582 MW; 3246EAA801B2A13F CRC64;

Query Match 18.6%; Score 21; DB 1; Length 22;
 Best Local Similarity 28.6%; Pred. No. 2.3e+03;
 Matches 2; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 Oy 11 LEEILMH 17
 Db 13 IEKVYVH 19

RESULT 13
 IAPP_LEPEU STANDARD; PRT; 23 AA.
 AC 007333;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE ISLET AMYLOID POLYPEPTIDE (AMYLIN) (FRAGMENT).
 GN IAPP.
 OS Lepus europaeus (European hare).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Lepus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PANCREAS;
 RX MEDLINE-93215963; PubMed-8462765;
 RA Christmanon L., Betsholtz C., Leckstroem A., Engstroem U.,
 RA Cortie C., Johnson K.H., Adrian T.E., Westermarck P.,
 RT "Islet amyloid polypeptide in the rabbit and European hare: studies
 RT on its relationship to amyloidogenesis.";
 RL Diabetologia 36:183-188(1993).
 CC -1- FUNCTION: SELECTIVELY INHIBITS INSULIN-STIMULATED GLUCOSE
 CC UTILIZATION AND GLYCOGEN DEPOSITION IN MUSCLE, WHILE NOT AFFECTING
 CC ADIPOCYTE GLUCOSE METABOLISM.
 CC -1- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: S57802; AAB26083.1; -
 CC DR Hormone; Amyloid.
 CC KW Hormone; Amyloid.
 CC FT NON_TER 1
 CC FT PEPTIDE <1 >23 ISLET AMYLOID POLYPEPTIDE.
 CC FT NON_TER 23
 CC SQ SEQUENCE 23 AA; 2546 MW; A5EE561D52B353DD CRC64;

Query Match 18.6%; Score 21; DB 1; Length 23;
 Best Local Similarity 44.4%; Pred. No. 2.4e+03;
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Oy 9 QNLEILMH 17
 Db 2 QRLANFLIH 10

RESULT 14
 RLS_HALHA STANDARD; PRT; 23 AA.
 ID RLS_HALHA
 AC P50356;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE 50S RIBOSOMAL PROTEIN L5P (HHA15) (FRAGMENT).
 GN RPL5P.
 OS Halobacterium halobium.
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
 OC Halobacterium.
 RN [1]
 RP SEQUENCE.
 RC STRAIN-DSM 670;
 RX MEDLINE-94229075; PubMed-8174557;
 RA McDougall J., Wilmann-Liebold B.,
 RT "Comparative analysis of the protein components from 5S rRNA, protein
 RT complexes of halophilic archaeobacteria.";
 RL Eur. J. Biochem. 221:779-785(1994).
 CC -1- SIMILARITY: BELONGS TO THE L5P FAMILY OF RIBOSOMAL PROTEINS.
 CC DR INTERPRO: IPR002132;
 DR PIR: C33084; C33084.
 DR PFAM: PF00281; RIBOSOMAL_L5; 1.
 DR PROSITE: PS00358; RIBOSOMAL_L5; PARTIAL.
 KW RIBOSOMAL protein.
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2773 MW; 19B318355A19B298 CRC64;

Query Match 18.6%; Score 21; DB 1; Length 23;
 Best Local Similarity 28.6%; Pred. No. 2.4e+03;
 Matches 2; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 Oy 11 LEEILMH 17
 Db 16 IEKVYVH 22

RESULT 15
 FKX7_PINPS STANDARD; PRT; 15 AA.
 ID FKX7_PINPS
 AC P81104;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 38, Last annotation update)
 DE 70 KDA PEPTIDYLPROLYL ISOMERASE (EC 5.2.1.8) (PEPTIDYLPROLYL CIS-TRANS
 DE ISOMERASE) (CYCLOPHILIN) (PIPIASE) (S1205-06) (FRAGMENT).
 OS Pinus pinaster (Maritime pine).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Coniferales; Coniferales; Pinaceae; Pinus.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-NEEDLE;
 RA Plomion C., Costa P., Bahman N., Frigerio J.-M.,
 RT "Genetic analysis of needle proteins in maritime pine. 1. Mapping
 RT dominant and codominant protein markers assayed on diploid tissue, in
 RT a haploid-based genetic map.";
 RL Silvae Genetica 46:161-165(1997).
 RN [2]
 RP SEQUENCE.
 RC TISSUE-NEEDLE;
 RX MEDLINE-99274088; PubMed-10344291;
 RA Costa P., Plomion C., Bauw G., Dubos C., Bahman N., Kremer A.,
 RA Frigerio J.-M., Plomion C.,
 RT "Separation and characterization of needle and xylem maritime pine
 RT proteins.";
 RL Electrophoresis 20:1098-1108(1999).
 CC -1- FUNCTION: PPIASE THAT BINDS CALMODULIN (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: PEPTIDYLPROLYL (OMEGA-180) -
 CC PEPTIDYLPROLYL (OMEGA-0).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
 CC 5.3. ITS MW IS: 72 KDA.
 CC -1- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.
 CC DR INTERPRO: IPR001179;
 DR PROSITE: PS00453; FKBP_PPIASE_1; PARTIAL.
 DR PROSITE: PS00454; FKBP_PPIASE_2; PARTIAL.
 DR PROSITE: PS00059; FKBP_PPIASE_3; PARTIAL.
 KW Isomerase; Rotamase; Repeat; Calmodulin-binding.
 FT NON_TER 1

FT NON_TER 15 15
SQ SEQUENCE 15 AA: 1675 MW: 2B53999722277E3F CRC64;

Query Match 17.7%: Score 20; DB 1; Length 15;
Best Local Similarity 25.0%: Pred. No. 2.2e+03;
Matches 3; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 NMELADOPONLE 12
: 11 : : 1
Db 4 SWETPETGEVE 15

RESULT 16
GLN2_PINPS STANDARD; PRT: 15 AA.
ID GLN2_PINPS
AC P81107;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE PROBABLE GLUTAMINE SYNTHETASE LEAF ISOZYME (EC 6.3.1.2) (GLUTAMATE--
AMMONIA LIGASE) (S2205/S2287) (N47/N48) (FRAGMENT).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Pinus.
RN [1]
RP SEQUENCE.
RC TISSUE=NEEDLE;
RA Plomion C., Costa P., Bahman N., Frigerio J.-M.;
RT "Genetic analysis of needle proteins in maritime pine. 1. Mapping
RT dominant and codominant protein markers assayed on diploid tissue, in
RT a haploid-based genetic map.";
RL Silvae Genetica 46:161-165(1997).
RN [2]
RP SEQUENCE.
RC TISSUE=NEEDLE;
RA MEDLINE=99274088; PubMed=10344291;
RX Costa P., Plomion C., Bauw G., Dubos C., Bahman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -1- FUNCTION: THE LIGHT-MODULATED CHLOROPLAST ENZYME, ENCODED BY A
CC NUCLEAR GENE AND EXPRESSED PRIMARILY IN LEAVES, IS RESPONSIBLE FOR
CC THE REASSIMILATION OF THE AMMONIA GENERATED BY PHOTORESPIRATION
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +
CC ORTHOPHOSPHATE.
CC -1- SUBUNIT: HOMOCOTYMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC 5.7, ITS MW IS: 42 KDA.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR INTERPRO: IPR001691;
DR PROSITE: PS00180; GMA_1; PARTIAL.
DR PROSITE: PS00181; GMA_ATP; PARTIAL.
KW Ligase; Chloroplast.
FT NON_TER 15 15
FT NON_TER 15 15
SQ SEQUENCE 15 AA: 1614 MW: 24A2420BEFDD60D27 CRC64;

Query Match 17.7%: Score 20; DB 1; Length 15;
Best Local Similarity 75.0%: Pred. No. 2.2e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMEL 4
: 11 : : 1
Db 3 NMPL 6

RESULT 17
VORA_METTM

ID VORA_METTM STANDARD; PRT: 15 AA.

AC P80907;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE KETOISOVALLERATE OXIDOREDUCTASE SUBUNIT VORA (EC 1.-.-.-) (VOR) (2-
DE OXISOVALLERATE OXIDOREDUCTASE ALPHA CHAIN) (2-OXISOVALLERATE-
DE FERREDOXIN OXIDOREDUCTASE ALPHA SUBUNIT) (FRAGMENT).
GN VORA.

OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.

RN [1]
RP SEQUENCE.
RX MEDLINE=97261844; PubMed=9108258;
RA Tersteegen A., Linder D., Thauer R.K., Hedderich R.;
RT "Structures and functions of four anabolic 2-oxoacid oxidoreductases
RT in Methanobacterium thermoautotrophicum.";
RL Eur. J. Biochem. 244:862-868(1997).
CC -1- SUBUNIT: HETEROTRIMER OF THE VORA, VORA AND VORC SUBUNITS.
CC -1- MISCELLANEOUS: THE PH OPTIMUM IS PH 9.7 AND THE OPTIMAL
CC TEMPERATURE IS 75 DEGREES CELSIUS.
KW Oxidoreductase.

FT NON_TER 15 15
SQ SEQUENCE 15 AA: 1779 MW: 3137086531CA528F CRC64;

Query Match 17.7%: Score 20; DB 1; Length 15;
Best Local Similarity 37.5%: Pred. No. 2.2e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 7 OPONIEEI 14
: 11 : : 1
Db 7 KPDSLEXV 14

RESULT 18
LE05_BIOGL STANDARD; PRT: 16 AA.
ID LE05_BIOGL
AC P80744;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEMOLYMPH 65 KDA LECTIN BG05 (FRAGMENT).
GN BG05.
OS Blomphalaria glabrata (Bloodfluke planorb).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Planorbidae; Blomphalaria.
RN [1]
RP SEQUENCE.
RX STRAIN=M-LINE; TISSUE=HEMOLYMPH;
RC MEDLINE=97385165; PubMed=9238039;
RA Adema C.M., Hertel L.A., Miller R.D., Loker E.S.;
RT "A family of fibrinogen-related proteins that precipitates parasite-
RT derived molecules is produced by an invertebrate after infection.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:8691-8696(1997).
CC -1- FUNCTION: BINDS AND PRECIPITATES ANTIGENS OF THE PARASITE
CC ECHINOSTOMA PARENSEI.
CC -1- SUBCELLULAR LOCATION: SECRETED IN THE HEMOLYMPH.
CC -1- INDUCTION: BY INFECTION.
KW Lectin.

FT NON_TER 16 16
FT NON_TER 16 16
SQ SEQUENCE 16 AA: 1790 MW: 57489A8F2EDDA94 CRC64;

Query Match 17.7%: Score 20; DB 1; Length 16;
Best Local Similarity 33.3%: Pred. No. 2.3e+03;
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 3 ELADOPONIEEI 14
: 11 : : 1
Db 2 ELADIAQYVDL 13

```

RESULT 19
EFG_THAQ          STANDARD:      PRT:      17 AA.
ID 001697:
AC 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DR 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR G (EF-G) (FRAGMENT).
GN FUS OR FUS.
OS Thermus aquaticus.
NC Bacteria; Thermus/Delinooccus group; Thermus group; Thermus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EP 00276:
RX MEDLINE-92362620: PubMed-1499561;
RA Voss R.H., Hartmann R.K., Lippmann C., Jahn O.,
RA Erdmann V.;
RT "Sequence of the tufA gene encoding elongation factor EF-Tu from
RT Thermus aquaticus and overproduction of the protein in Escherichia
RT coli.";
RL Eur. J. Biochem. 207:839-846(1992).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC RIBOSOME.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC DR EMBL: X63432; CAA46997.1;
CC DR PIR: S29294; S29294.
CC DR HSSP: P13511; LEO.
CC DR INTERPRO: IPR000795;
CC DR PROSITE: PS00301; EFACITOR-GTP; PARTIAL.
CC Elongation factor; Protein biosynthesis; GTP-binding.
CC NON_TER 1
FT SEQUENCE 17 AA; 2094 MM; EA46E1EF05F86E1D CRC64;
SQ

```

```

Query Match          17.7%; Score 20; DB 1; Length 17;
Best Local Similarity 33.3%; Pred. No. 2.5e+03;
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 8 PONTREILHQC 19
ID 1: 1: 1: 1: 1:
DB 6 PROIOELIKQ 17

RESULT 20
DCAM_ACACA          STANDARD:      PRT:      19 AA.
ID AC P34039:
AC 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DR 01-FEB-1994 (Rel. 28, Last annotation update)
DE S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA SUBUNIT (BC 4.1.1.50)
DE (ADOMETDC) (FRAGMENT).
OS Acanthamoeba castellanii (Amoeba).
NC Eukaryota; Acanthamoebidae; Acanthamoeba.
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE-94029912; PubMed-8216217;
RA Hugo E.R., Byers T.J.;
RT "S-adenosyl-L-methionine decarboxylase of Acanthamoeba castellanii

```

```

RT (Neff): purification and properties.";
RL Biochem. J. 295:203-209(1993).
CC -1- FUNCTION: S-ADENOSYLMETHIONINE DECARBOXYLASE IS ESSENTIAL FOR THE
CC BIOSYNTHESIS OF SPERMINE AND SPERMIDINE. THE ALPHA SUBUNIT
CC CONTAINS THE ACTIVE SITE.
CC -1- CATALYTIC ACTIVITY: S-ADENOSYLMETHIONINE - (5-DEOXY-5-ADENOSYL)
CC (3-AMINOPROPYL)METHYLSULFONIDIUM SALT + CO(2).
CC -1- COFACTOR: REQUIRES A PYRUVYL GROUP FOR ITS ACTIVITY.
CC -1- PATHWAY: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE
CC AMINOPROPYL MOIETY REQUIRED FOR SPERMIDINE AND SPERMINE
CC BIOSYNTHESIS FROM PUTRESCINE.
CC -1- SUBUNIT: HETEROETRAMER OF TWO ALPHA AND TWO BETA CHAINS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ONLY DURING EXPONENTIAL GROWTH.
CC -1- INDUCTION: STIMULATED BY PUTRESCINE. INHIBITED BY AROMATIC
CC DIAMIDINES BERENIL, PENTAMIDINE, PROPAMIDINE, HYDROXYSTILBAMIDINE,
CC BY ETHIDIUM BROMIDE AND METHYLGLOXAL.
CC -1- SIMILARITY: BELONGS TO THE ADOMETC FAMILY.
CC INTERPRO: IPR001985;
CC DR PROSITE: PS01361; ADOMETDC; PARTIAL.
CC KW Spermidine biosynthesis; Lyase; Decarboxylase; Pyruvate.
CC FT MOD_RES 1 1
CC FT NON_TER 1 1
CC FT SEQUENCE 19 AA; 2153 MM; 88B18AD9B6142AEF CRC64;
SQ

```

```

Query Match          17.7%; Score 20; DB 1; Length 19;
Best Local Similarity 45.5%; Pred. No. 2.8e+03;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 10 NIEELIMHCOT 20
ID 1: 1: 1: 1: 1:
DB 7 NKLRLIKGT 17

```

```

RESULT 21
EHPA_CLOPA          STANDARD:      PRT:      19 AA.
ID EHPA_CLOPA
AC P81342;
DR 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DR 15-JUL-1998 (Rel. 36, Last annotation update)
DE ELECTRON TRANSFER FLAVOPROTEIN ALPHA-SUBUNIT (ALPHA-ETF) (ELECTRON
DE TRANSFER FLAVOPROTEIN LARGE SUBUNIT) (ETFLS) (CP 14) (FRAGMENT).
GN ETFA.
OS Clostridium pasteurianum.
NC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
CC Clostridium.
CC [1]
CC RP SEQUENCE.
CC RC STRAIN-W5;
CC RX MEDLINE-98291870; PubMed-9629918;
CC RA Flensburg R., Skjeldal L.;
CC RT "Two-dimensional gel electrophoresis separation and N-terminal
CC sequence analysis of proteins from Clostridium pasteurianum W5.";
CC RL Electrophoresis 19:802-806(1998)
CC -1- FUNCTION: THE ELECTRON TRANSFER FLAVOPROTEIN SERVES AS A SPECIFIC
CC ELECTRON ACCEPTOR FOR OTHER DEHYDROGENASES. IT TRANSFERS THE
CC OXIDOREDUCTASE (ETF DEHYDROGENASE) (BY SIMILARITY).
CC -1- COFACTOR: CONTAINS ONE MOLECULE OF FAD PER DIMER (BY SIMILARITY).
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE ETF ALPHA-SUBUNIT / FIXB FAMILY.
CC DR INTERPRO: IPR001308;
CC DR PROSITE: PS00696; ETF_ALPHA; PARTIAL.
CC KW Electron transport; Flavoprotein; FAD.
CC FT NON_TER 19
CC FT SEQUENCE 19 AA; 2110 MM; BB23312B28F4F0D6 CRC64;
SQ

```

```

Query Match          17.7%; Score 20; DB 1; Length 19;
Best Local Similarity 50.0%; Pred. No. 2.8e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

OY 2 WELADQ 7
1 1 1
DB 11 WVFADQ 16

RESULT 22

OXLA_OPHHA STANDARD; PRT; 19 AA.
ID OXLA_OPHHA
AC P81383;
RT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
OS L-AMINO ACID OXIDASE (EC 1.4.3.2) (LMO) (FRAGMENT).
OC Ophiophagus hannah (King cobra) (Naja hannah).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubridae;
OC Elapidae; Elapinae; Ophiophagus.
RN [1]

SEQUENCE.

RP TISSUE-VENOM;
RX MEDLINE-94361525; PubMed-8080286;
RA Pennudurai G., Chung M.C.M., Tan N.-H.;
RT "Purification and properties of the L-amino acid oxidase from Malayan
pit viper (Calloselasma rhodostoma) venom.";
RL Arch. Biochem. Biophys. 313:373-378(1994).
RN [2]

SEQUENCE OF 1-15.

RP TISSUE-VENOM;
RX MEDLINE-97449790; PubMed-9304806;
RA Ahn M.Y., Lee B.M., Kim Y.S.;
RT "Characterization and cytotoxicity of L-amino acid oxidase from the
venom of King cobra (Ophiophagus hannah).";
RL Int. J. Biochem. Cell Biol. 29:911-919(1997).
CC -I- FUNCTION: HAS CYTOTOXIC ACTIVITY (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: A L-AMINO ACID + H(2)O + O(2) -> A 2-OKO ACID +
NH(3) + H(2)O(2).
CC -I- COFACTOR: FAD.
CC -I- SUBUNIT: HOMODIMER (PROBABLE).
CC -I- PFM: GLYCOSTYLATED.
CC -I- SIMILARITY: BELONGS TO THE FLAVIN MONOXAMINE OXIDASE FAMILY.
CC STRONG, TO MOUSE FIG-1.
KM Oxidoreductase; Flavoprotein; FAD; Glycoprotein; Venom.
FT CONFLICT 1 1 H -> S (IN REF. 2).
FT NON_TER 19 19
SQ SEQUENCE 19 AA: 2298 MW; DD911A5B414F1427 CRC64;

Query Match 17.7%; Score 20; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 NLEE 13
1 1 1 1 1

DB 4 NLEE 7

RESULT 23

DFTS_RAT STANDARD; PRT; 20 AA.
ID DFTS_RAT
AC P07448;
RT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DE 01-MAR-1989 (Rel. 10, Last annotation update)
OS DENTINAL FLUID TRANSPORT-STIMULATING PEPTIDE (DFT-STIMULATING
PEPTIDE).
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
RN [1]

RP TISSUE-PAROTID GLAND;
RX MEDLINE-87131231; PubMed-3815601;
RA Yamamoto T., Kobayashi M., Kobayashi M., Yamamoto M., Nomura M.,

RA Aonuma S.;
RT "Isolation and amino acid sequence of dentinal fluid transport-
stimulating peptide from rat parotid glands.";
RL Chem. Pharm. Bull. 34:3803-3811(1986).
RN [2]

CHARACTERIZATION.

RX MEDLINE-67131708; PubMed-5297832;
RA Steinman R.R.;
RT "The movement of acriflavine hydrochloride through molars of rats on
a cariogenic and non-cariogenic diet.";
RL J. South. Calif. Dent. Assoc. 35:151-157(1967).
CC -I- FUNCTION: THIS PEPTIDE STIMULATES THE TRANSPORT OF DENTINAL FLUID,
WHICH IS IMPORTANT FOR THE PREVENTION OF DENTAL CARIES.
DR PIR; J00001; DIRT.
KM Dental caries; Parotid gland; Hormone.
SQ SEQUENCE 20 AA: 2165 MW; FA164F2B6AF80D5A CRC64;

Query Match 17.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 3e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 WEL 4
1 1 1
DB 5 WEL 7

RESULT 24

PSAL_SYNVU STANDARD; PRT; 20 AA.
ID PSAL_SYNVU
AC P25937;
RT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
OS PHOTOSYSTEM I REACTION CENTRE SUBUNIT XI (FRAGMENT).
GN PSAL.
OC Synecococcus vulcanus.
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
RN [1]

RP SEQUENCE.
RX MEDLINE-89338747; PubMed-2503399;
RA Kelle H., Ikeuchi M., Hiyaama T., Inoue Y.;
RT "Identification of photosystem I components from the cyanobacterium,
Synecococcus vulcanus by N-terminal sequencing.";
RL FEBS Lett. 253:257-263(1989).
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -I- SIMILARITY: BELONGS TO THE PSAL FAMILY.
DR PIR; S05320; S05220.
KM Photosystem I; Photosynthesis; Transmembrane.
FT NON_TER 20 20
SQ SEQUENCE 20 AA: 2170 MW; 730FECDD2EA02A2C CRC64;

Query Match 17.7%; Score 20; DB 1; Length 20;
Best Local Similarity 28.6%; Pred. No. 3e+03;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 7 QPONEEILMHCO 20
1 1 1 1 1 1 1
DB 6 KPYNGDPFVGHLS 19

RESULT 25

RAN_XENLA STANDARD; PRT; 24 AA.
ID RAN_XENLA
AC P52301;
RT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
OS GTP-BINDING NUCLEAR PROTEIN RAN (TC4) (FRAGMENTS).
GN RAN.
OC Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia: Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE.
RC TISSUE-OVARY;
RX MEDLINE-94019818; PubMed-8413630;
RA Moore M.S., Blobel G.;
RT "The GTP-binding protein Ran/TCA is required for protein import into
the nucleus.";
RL Nature 365:661-663(1993).
CC -1- FUNCTION: GTP-BINDING PROTEIN INVOLVED IN NUCLEOCYTOPLASMIC
TRANSPORT. REQUIRED FOR THE IMPORT OF PROTEIN INTO THE NUCLEUS AND
ALSO FOR RNA EXPORT. INVOLVED IN CHROMATIN CONDENSATION AND
CONTROL OF CELL CYCLE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE RAN FAMILY IN THE RAS SUPERFAMILY.
DR HSP: P28746; 1A2K.
DR INTERPRO: IPR002041; -
DR PROSITE; PS01115; RAN; PARTIAL.
KW GTP-binding; Nuclear protein; Protein transport.
FT NON_TER 1 1
FT NON_CONS 13 14 GTP (BY SIMILARITY).
FT NP_BIND 18 22
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2657 MW; B69F83236247A250 CRC64;

Query Match

17.7%; Score 20; DB 1; Length 24;

Best Local Similarity 50.0%; Pred. No. 3.6e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 WELADQ 7
| : | |
Db 17 WDRAGQ 22

Search completed: February 5, 2001, 10:56:00
Job time: 500 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:55:22; Search time 93.77 Seconds
(without alignments)
24.999 Million cell updates/sec

Title: US-08-981-824-3

Perfect score: 113
Sequence: 1 NMEIADQPONLEIIMHCQT 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database:

SPTREMBL.15:*
1: SP_Archaea:*
2: SP_Bacteria:*
3: SP_Fungi:*
4: SP_Human:*
5: SP_Invertebrate:*
6: SP_Mammal:*
7: SP_Mhc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	29.2	18	12	078376 human immun
2	33	29.2	23	8	09MG67 pinus ponde
3	29	23.7	17	5	09TWB9 acanthamoeb
4	29	25.7	20	2	09RAW5 helicobacte
5	29	25.7	20	3	09URC7 saccharomyc
6	29	25.7	22	12	084254 bovine papil
7	28	24.8	18	2	09RSF6 helicobacte
8	27	23.9	18	2	047137 escherichia
9	27	23.9	24	10	09S883 chlamydomon
10	26	23.0	16	5	09TMC0 acanthamoeb
11	26	23.0	18	6	09TRQ2 bos taurus
12	26	23.0	20	13	09PS25 rana catesb
13	26	23.0	23	2	09R313 chlamydia t
14	25	22.1	16	6	09TRB4 bos taurus
15	25	22.1	19	2	09RAU6 lactobacilli
16	25	22.1	24	12	09PXB7 hog cholera
17	25	22.1	25	4	09UOB1 Oryzobolus
18	24	21.2	14	11	09R1G8 rattus norv
19	24	21.2	14	12	010234 human immun

20	24	21.2	17	7	030218 homo sapien
21	24	21.2	21	4	09UC77 homo sapien
22	24	21.2	23	11	097918 mus musculu
23	24	21.2	25	4	016308 homo sapien
24	23.5	20.8	24	4	09UCG1 homo sapien
25	23.5	20.8	24	6	09TRP6 sus scrofa
26	23	20.4	15	10	09T331 pisum sativ
27	23	20.4	19	2	09K4X0 Oryzobolus
28	23	20.4	20	10	09S8Y7 raphanus sa
29	23	20.4	21	4	09UC16 homo sapien
30	23	20.4	22	2	052009 pseudomonas
31	23	20.4	22	2	09KK60 mycobacteri
32	23	20.4	24	4	016476 homo sapien
33	23	20.4	24	5	094373 caenorhabdi
34	23	20.4	24	10	09S8H6 Oryzobolus
35	22	19.5	15	5	026159 escherichia
36	22	19.5	15	5	026159 plasmodium
37	22	19.5	17	5	09NFC3 dirosophila
38	22	19.5	20	2	097160 salmoneilla
39	22	19.5	21	1	09UWH2 thermococcu
40	22	19.5	21	2	056354 paracoccu
41	22	19.5	21	4	09UC33 homo sapien
42	22	19.5	21	5	09TWM1 lytechinus
43	22	19.5	21	11	063076 rattus norv
44	22	19.5	23	2	09S899 chlamydia t
45	22	19.5	24	2	055238 salmoneilla

ALIGNMENTS

RESULT 1
ID 078376 PRELIMINARY: PRT; 18 AA.
AC 078376;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 08, Last annotation update)
DE VIRAL SAMPLE FLPR41 (FLORIDA PATIENT B), PARTIAL ENV CDS, V4 REGION
DE (FRAGMENT).
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang L.O., Leigh-Brown A.J.;
RA Submitted (Apr-1992) to the EMBL/Genbank/DBJ databases.
RN [2]
RX MEDLINE=92271245; PubMed=1589796;
RA Ou C.Y., Ciesielski C.A., Myers G., Banda C.I., Luo C.C.,
RA Korber B.T.M., Mullins J.I., Schochetman G., Berkelman R.L.,
RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
RA Curran J.W., Jaffe H.W.;
RT "Molecular epidemiology of HIV transmission in a dental practice";
RT Science 256:1165-1171(1992).
DR EMBL; M92122; AAA44492.1;
FT NON_TER 1
FT 1
SQ SEQUENCE 18 AA; 2000 MW; 0BIE4794679E050A CRC64;

Query Match 29.2%; Score 33; DB 12; Length 18;
Best Local Similarity 41.7%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 8 PONTLEIIMHCQ 19
DB 7 PNTLEIIMHCQ 18
PONTLEIIMHCQ 19
PNTLEIIMHCQ 18

RESULT 2
OY 8 PONTLEIIMHCQ 19
DB 7 PNTLEIIMHCQ 18
PONTLEIIMHCQ 19
PNTLEIIMHCQ 18

ID Q9MG67 PRELIMINARY; PRT; 23 AA.
 AC Q9MG67;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE NAH DEHYDROGENASE SUBUNIT 1 (FRAGMENT).
 GN NAD1.
 OS Pinus ponderosa.
 OC Mitochondrion.
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Coniferopsida; Coniferales; Pinaceae; Pinus.
 CX NCBI_TaxID=55062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kreiser B.R., Milton J.B., Rehfeldt G.E.;
 RT "Primers designed to amplify a mitochondrial NAD1 intron in ponderosa
 pine (Pinus ponderosa), lamber pine (P. flexilis), and Scots pine (P.
 sylvestris)."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF231325; AAF78900.1;
 KW Mitochondrion.
 FT NON_TER 1 1
 FT SEQUENCE 23 AA; 2614 MW; 52FAB49CF962C2DC CRC64;
 SO SEQUENCE

Query Match 29.2%; Score 33; DB 8; Length 23;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 8 PONEEILM 16
 Db 7 PNLSEIIV 15

RESULT 3
 O9TMB9 PRELIMINARY; PRT; 17 AA.
 AC Q9TMB9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE PROFILIN-BINDING CORTICAL COMPLEX-18 KDA POLYPEPTIDE.
 OS Acanthamoeba castellanii (Amoeba).
 CC Eukaryota; Acanthamoebidae; Acanthamoeba.
 CX NCBI_TaxID=5755;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95014701; PubMed=7929556;
 RA Machesky L.M., Atkinson S.J., Ampe C., Vandekerckhove J.,
 RA Pollard T.D.;
 RT "Purification of a cortical complex containing two unconventional
 RT actins from Acanthamoeba by affinity chromatography on profilin-
 RT agarose."
 RL J Cell Biol. 127:107-115(1994).
 SO SEQUENCE 17 AA; 1875 MW; 7260445CDF5AE78B CRC64;

Query Match 25.7%; Score 29; DB 5; Length 17;
 Best Local Similarity 45.5%; Pred. No. 4.6e+02;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 5 ADOPONEEIL 15
 Db 5 ADQPDIDEAI 15

RESULT 4
 O9R4W5 PRELIMINARY; PRT; 20 AA.
 AC Q9R4W5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE 30 KDA MAJOR HEAT SHOCK PROTEIN (FRAGMENT).
 OS Helicobacter pylori (Campylobacter pylori).
 CC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 CC Helicobacter.
 CX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95020803; PubMed=7935068;
 RA Yokota K., Hirai Y., Hague M., Hayashi S., Isogai H., Sugiyama T.,
 RA Nagamachi E., Tsukada Y., Fujii N., Oguma K.;
 RT "Heat shock protein produced by Helicobacter pylori."
 RL Microbiol. Immunol. 38:405-405(1994).
 DR INTERPRO: IPR002026;
 DR PFM: PFO0547; urease-gamma; 1.
 SO SEQUENCE 20 AA; 2302 MW; 29C9DFBFD6D21805 CRC64;

Query Match 25.7%; Score 29; DB 2; Length 20;
 Best Local Similarity 30.0%; Pred. No. 5.3e+02;
 Matches 3; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 8 PONEEILMH 17
 Db 5 PKEIDKMLH 14

RESULT 5
 O9URC7 PRELIMINARY; PRT; 20 AA.
 AC Q9URC7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE LIPID-BINDING PROTEIN.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=91353077; PubMed=1882548;
 RA Creutz C.E., Snyder S.L., Kambouris N.G.;
 RT "Calcium-dependent secretory vesicle-binding and lipid-binding
 RT proteins of Saccharomyces cerevisiae."
 RL Yeast 7:229-244(1991).
 SO SEQUENCE 20 AA; 2388 MW; 594377C8C3E72B0D CRC64;

Query Match 25.7%; Score 29; DB 3; Length 20;
 Best Local Similarity 42.9%; Pred. No. 5.3e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

OY 2 WELADOPONEEIL 15
 Db 5 WD--DDEINDEL 16

RESULT 6
 O84254 PRELIMINARY; PRT; 22 AA.
 AC O84254;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE X PROTEIN (FRAGMENT).
 OS Bovine papillomavirus.
 CC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
 CX NCBI_TaxID=10571;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89067912; PubMed=2848926;
 RA Stamps A.C., Campo M.S.;

"Mapping of two novel transcripts of bovine papillomavirus type 4.";
J. Gen. Virol. 69:3033-3045(1988).

DR EMBL; M35264; AAA46926.1;
FT NON_TER 1 1
FT NON_TER 22 22
SO SEQUENCE 22 AA: 2273 MW: 2801BC23480C9CF9 CRC64;

Query Match

Best Local Similarity 25.7%; Score 29; DB 12; Length 22;
Best Local Similarity 45.5%; Pred. No. 6e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 ELADOPONLE 13
DB 12 DVADRPDLPE 22

RESULT 7

O9R5F6 PRELIMINARY: PRT: 18 AA.
AC O9R5F6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE UREASE SMALL SUBUNIT (FRAGMENT).
OS Helicobacter mustelae.
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCB1_TaxID=217;
RN [1]
RP SEQUENCE.
RX MEDLINE-93084378; PubMed-1452359;
RA Turbett G.R., Hoj P.B., Horne R., Mee B.J.;
RT "Purification and characterization of the urease enzymes of
RT Helicobacter species from humans and animals."
RL Infect. Immun. 60:5259-5266(1992).
SQ SEQUENCE 18 AA: 2060 MW: 29C8E6AB7E21805 CRC64;

Query Match

Best Local Similarity 24.8%; Score 28; DB 2; Length 18;
Best Local Similarity 30.0%; Pred. No. 7.1e+02;
Matches 3; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 8 PONEELIMH 17
DB 3 PKELDKMLH 12

RESULT 8

ID O47137 PRELIMINARY: PRT: 18 AA.
AC O47137;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE POT. URF.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCB1_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Bouche J.P.;
RL Submitted (Apr-1988) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-88289404; PubMed-3041373;
RA Cam K., Bejar S., Gil D., Bouche J.P.;
RT "Identification and sequence of gene dicB: translation of the division
RT inhibitor from an in-phase internal start."
RT Nucleic Acids Res. 16:6327-6338(1988).
RN [3]
RP SEQUENCE FROM N.A.

RX MEDLINE-88232418; PubMed-2836697;
RA Bejar S., Bouche F., Bouche J.P.;
RT "Cell division inhibition gene dicB is regulated by a locus similar to
RT lambdaoid bacteriophage immunity loci."
RL Mol. Gen. Genet. 212:11-19(1988).
RN [4]
RP SEQUENCE FROM N.A.
RA Bouche J.P.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE-87016327; PubMed-3332030;
RA Bejar S., Cam K., Bouche J.P.;
RT "Control of cell division in Escherichia coli. DNA sequence of dicA
RT and of a second gene complementing mutation dicA1, dicC."
RL Nucleic Acids Res. 14:6821-6833(1986).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE-90014174; PubMed-2477663;
RA Bouche F., Bouche J.P.;
RT "Genetic evidence that DicF, a second division inhibitor encoded by
RT the Escherichia coli dicB operon, is probably RNA."
RN Mol. Microbiol. 3:991-994(1989).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE-90219018; PubMed-1691289;
RA Faubladier M., Cam K., Bouche J.P.;
RT "Escherichia coli cell division inhibitor DicF-RNA of the dicB operon.
RT Evidence for its generation in vivo by transcription termination and
RT by RNase III and RNase E-dependent processing."
RL J. Mol. Biol. 212:461-471(1990).
DR EMBL; X07465; CAA30350.1;
SQ SEQUENCE 18 AA: 2114 MW: 28B470E8012EC7A2 CRC64;

Query Match

Best Local Similarity 23.9%; Score 27; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 MHCO 19
DB 1 MHCO 4

RESULT 9

ID O9S883 PRELIMINARY: PRT: 24 AA.
AC O9S883;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE ATP SYNTHASE DELTA SUBUNIT (FRAGMENT).
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCB1_TaxID=3055;
RN [1]
RP SEQUENCE.
RX MEDLINE-96128220; PubMed-8543042;
RA Fiedler H.R., Schmid R., Ieu S., Shavitt N., Strothmann H.;
RT "Isolation of CFocFI from Chlamydomonas reinhardtii cw15 and the N-
RT terminal amino acid sequences of the CFocFI subunits."
RT FEBS Lett. 377:165-166(1995).
SQ SEQUENCE 24 AA: 2707 MW: 50CAE3310F50FB44 CRC64;

Query Match

Best Local Similarity 23.9%; Score 27; DB 10; Length 24;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 ELADOPONLE 12
DB 15 ELADKGRLE 24

```

RESULT 10
O9TRC0 ID Q9TRC0 PRELIMINARY; PRT; 16 AA.
AC Q9TRC0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE ACTIN (FRAGMENT).
OS Acanthamoeba castellanii (Amoebea).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_Taxid=5755;
RN [1]
RP SEQUENCE.
RX MEDLINE=95014701; PubMed=7929556;
RA Machesky L.M., Atkinson S.J., Ampe C., Vandekerckhove J.,
RA Pollard T.D.;
RT "Purification of a cortical complex containing two unconventional
RT actins from Acanthamoeba by affinity chromatography on profilin-
RT agarose."
RL J. Cell Biol. 127:107-115(1994).
SQ SEQUENCE 16 AA; 2115 MW; A64E24880BA06C4 CRC64;

Query Match 23.0%; Score 26; DB 5; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NWELADOPON 10
Db 1111111111
1 NWELADOPON 10

RESULT 11
O9TR02 ID Q9TR02 PRELIMINARY; PRT; 18 AA.
AC Q9TR02;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE P68 KINASE INHIBITOR (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=92332534; PubMed=1378438;
RA Lee T.G., Tomita J., Hovanessian A.G., Katze M.G.;
RT "Characterization and regulation of the 58,000-dalton cellular
RT inhibitor of the interferon-induced, dsRNA-activated protein kinase."
RL J. Biol. Chem. 267:14238-14243(1992).
SQ SEQUENCE 18 AA; 2206 MW; 91778DDE6B34082D CRC64;

Query Match 23.0%; Score 26; DB 6; Length 18;
Best Local Similarity 38.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 WELADOPONLEI 14
Db 1111111111
2 WELADOPONLEI 14
6 YELGDHELRLREV 18

RESULT 12
O9PS25 ID Q9PS25 PRELIMINARY; PRT; 20 AA.
AC Q9PS25;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

```

```

DE MONOMERIC ALPHA-MACROGLOBULIN PROTEINASE INHIBITOR (FRAGMENT).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_Taxid=8400;
RN [1]
RP SEQUENCE.
RX MEDLINE=93176138; PubMed=7679897;
RA Rubenstein D.S., Thogersen I.B., Pizzo S.V., Englund J.J.;
RT "Identification of monomeric alpha-macroglobulin proteinase inhibitors
RT in birds, reptiles, amphibians and mammals, and purification and
RT characterization of a monomeric alpha-macroglobulin proteinase
RT inhibitor from the American bullfrog Rana catesbeiana."
RL Biochem. J. 290:85-95(1993).
SQ SEQUENCE 20 AA; 2332 MW; 7CE78C52053F74D7 CRC64;

Query Match 23.0%; Score 26; DB 13; Length 20;
Best Local Similarity 83.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 PONTLEE 13
Db 1111111111
10 PONTLEE 15

RESULT 13
O9R313 ID Q9R313 PRELIMINARY; PRT; 23 AA.
AC Q9R313;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE VIRULENCE PROTEIN PG3-D (FRAGMENT).
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_Taxid=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L2 434B;
RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT "Gene identification of Chlamydia trachomatis by random DNA
RT sequencing."
RL Submitted (MUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087345; AAD04118.1;
DR EMBL; AF087324; AAD04099.1;
KW Hypothetical protein.
FT NON TER 1 23
FT NON TER 1 23
SQ SEQUENCE 23 AA; 2475 MW; BCBCE18F08862FE8B CRC64;

Query Match 23.0%; Score 26; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 PONTLEE 15
Db 1111111111
9 PONTLEE 16

RESULT 14
O9TRB4 ID Q9TRB4 PRELIMINARY; PRT; 16 AA.
AC Q9TRB4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE ATP-DEPENDENT 20 S PROTEASOME ACTIVATOR (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.

```

OX NCBI_TaxID-9913;
 RN [1]
 RP SEQUENCE:
 RX MEDLINE-94342244; PubMed-8063704;
 RA Demattio G.N., Moormaw C.R., Zagnitko O.P., Proske R.J., Chu-Ping M.,
 AFendis S.J., Swaffield J.C., Slaughter C.A.:
 RT "PA700, an ATP-dependent activator of the 20 S proteasome, is an
 RT ATPase containing multiple members of a nucleotide-binding protein
 RT family."
 RL J. Biol. Chem. 269:20878-20884(1994).
 SO SEQUENCE 16 AA; 1878 MW; F70F74211BE26EDE CRC64;

Query Match 22.1%; Score 25; DB 6; Length 16;
 Best Local Similarity 50.0%; Pred. No. 1.9e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 1 NWEADOP 8
 Db 7 NFOLDNP 14

RESULT 15
 O9R4U6 PRELIMINARY; PRT; 19 AA.
 AC O9R4U6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE BACTERIOCIN LACTACIN B INDUCER (FRAGMENT).
 OS Lactobacillus acidophilus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID-1579;
 RN [1]
 RP SEQUENCE:
 RX MEDLINE-95077353; PubMed-7986029;
 RA Barefoot S.F., Chen Y.R., Hughes T.A., Bodine A.B., Shearer M.Y.,
 RA Hughes M.D.:
 RT "Identification and purification of a protein that induces production
 RT of the Lactobacillus acidophilus bacteriocin lactacin B."
 RL Appl. Environ. Microbiol. 60:3522-3528(1994).
 SO SEQUENCE 19 AA; 2171 MW; 3221F0603F19E32F CRC64;

Query Match 22.1%; Score 25; DB 2; Length 19;
 Best Local Similarity 44.4%; Pred. No. 2.2e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 NWEADOPQ 9
 Db 9 NWKLMNPK 17
 RESULT 16
 O9PXB7 PRELIMINARY; PRT; 24 AA.
 AC O9PXB7;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE GLYCOPROTEIN E0 (FRAGMENT).
 OS Hog cholera virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Pestivirus.
 OX NCBI_TaxID-11096;
 RN [1]
 RP SEQUENCE:
 RX MEDLINE-93267778; PubMed-8388499;
 RA Rumenapf T., Unger G., Strauss J.H., Thiel H.J.:
 RT "Processing of the envelope glycoproteins of pestiviruses."
 RL J. Virol. 67:3288-3294(1993).
 SO SEQUENCE 24 AA; 2752 MW; A48D322C89550658 CRC64;

Query Match 22.1%; Score 25; DB 12; Length 24;
 Best Local Similarity 44.4%; Pred. No. 2.8e+03;
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 OY 2 WELADOPQN 10
 Db 6 WNLSDNGTN 14

RESULT 17
 O9UOB1 PRELIMINARY; PRT; 25 AA.
 AC O9UOB1;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE CD22 PROTEIN (FRAGMENT).
 GN CD22.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hattia Y., Tsuchiya N., Matsushita M., Shiota M., Hagihara K.,
 RA Tokunaga K.:
 RT "Identification of the gene variations in human CD22."
 DR EMBL: AB013005; BAA36574.1; -.
 FT NON_TER 1
 FT NON_TER 25
 SO SEQUENCE 25 AA; 2969 MW; 354944648D268289 CRC64;

Query Match 22.1%; Score 25; DB 4; Length 25;
 Best Local Similarity 33.3%; Pred. No. 3e+03;
 Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 WELADOPNLEE 13
 Db 1 WKRTOSQGLQE 12
 RESULT 18
 O9RIG8 PRELIMINARY; PRT; 14 AA.
 AC O9RIG8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE INSULIN RECEPTOR PRECURSOR (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID-10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY;
 RA Liu Y., Tam J.W.O.:
 RT "Partial sequence of rat insulin receptor gene."
 RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF110222; AAD40897.1; -.
 DR EMBL: AF110221; AAD40897.1; JOINED.
 KW Receptor.
 FT NON_TER 1
 FT NON_TER 14
 SO SEQUENCE 14 AA; 1826 MW; 51C2994579D01697 CRC64;

Query Match 21.2%; Score 24; DB 11; Length 14;
 Best Local Similarity 27.3%; Pred. No. 2.4e+03;

Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 8 PONTLEIIMHC 18
1 PERLTDLIMRC 11

RESULT 19

ID 010234 PRELIMINARY; PRT; 14 AA.
AC 010234;
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DE 01-JUL-2000 (TREMblrel. 13, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Homo immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCB1_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98216723; PubMed=9557645;
RA Salvi R., Garbuglia A.R., Di Caro A., Pulciani S., Montella F.,
Benedetto A.;
RT "Grossly defective nef gene sequences in a human immunodeficiency
virus type 1 seropositive long-term nonprogressor.";
RL J. Virol. 72:3646-3657(1998).
DR EMBL; U89854; AAC26093.1; -.
DR INTERPRO: IPR000328; -.
DR PFAM: PF00517; GP41; 1.
KW Envelope protein.
FT NON-TER 1
SQ SEQUENCE 14 AA; 173 MW; D5E74A99D45D0566 CRC64;

Query Match 21.2%; Score 24; DB 12; Length 14;
Best Local Similarity 62.5%; Pred. No. 2.4e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 9 ONLEIIM 16
1 1 1 1 1
7 OGIERILL 14

RESULT 20

ID 030218 PRELIMINARY; PRT; 17 AA.
AC 030218;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-JUN-2000 (TREMblrel. 14, Last annotation update)
DE FCER2 PROTEIN (FRAGMENT).
GN FCER2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96124133; PubMed=8552454;
RA Forster H.H., Masch R., Kreschmar T., Mischke D.,
Uchanska-Ziegler B., Ziegler A., Schmitt M., Wahn H.U.;
RT "Genetic markers on chromosome 19p and prenatal diagnosis of HLA class
II-deficient combined immunodeficiency.";
RL Pediatr. Res. 38:812-816(1995).
DR EMBL; S81114; AAB35925.1; -.
KW MHC.
FT NON-TER 1
SQ SEQUENCE 17 AA; 2076 MW; 53F5D4E75F1E5F47 CRC64;

Query Match 21.2%; Score 24; DB 7; Length 17;
Best Local Similarity 33.3%; Pred. No. 2.9e+03;

Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 WEIADOPONLEE 13
1 WDTQSLKQLEE 12

RESULT 21

ID 09UCH7 PRELIMINARY; PRT; 21 AA.
AC 09UCH7;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE TN GLYCOPHORIN A-TN ANTIGEN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93211988; PubMed=7681597;
RA Nakada H., Inoue M., Numata Y., Tanaka N., Funakoshi I., Fukui S.,
Mellors A., Yamashina I.;
RT "Epitopic structure of Tn glycophorin A for an anti-Tn antibody (MLS
128).";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2495-2499(1993).
SQ SEQUENCE 21 AA; 2230 MW; 29F2DD83C77FCE2 CRC64;

Query Match 21.2%; Score 24; DB 4; Length 21;
Best Local Similarity 50.0%; Pred. No. 3.6e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 13 EYIMHCOT 20
1 1 1 1 1
5 EVAMHTST 12

RESULT 22

ID P97918 PRELIMINARY; PRT; 23 AA.
AC P97918;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DE 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE FC GAMMA RECEPTOR (FRAGMENT).
GN FCGR2B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE OF 226-248 FROM N.A.
RX MEDLINE=8631694; PubMed=2944118;
RA Hibbs M.L., Walker I.D., Kirschaum L., Peltersz G.A., Deacon N.J.,
Chambers G.W., McKenzie I.F.C., Hogarth P.M.;
RT "The murine Fc receptor for immunoglobulin: purification, partial
amino acid sequence, and isolation of cDNA clones.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:6980-6984(1986).
DR EMBL; M14277; AAA37606.1; -.
DR MGD; MGI:95499; Fcgr2b.
FT NON-TER 1
SQ SEQUENCE 23 AA; 2670 MW; 13303A118E855AA6 CRC64;

Query Match 21.2%; Score 24; DB 11; Length 23;
Best Local Similarity 31.2%; Pred. No. 3.9e+03;
Matches 5; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 2 WEIADOPONTLEIIMH 17
1 1 1 1 1

Db 8 YSLKHPALDEETEH 23

RESULT 23

ID 016308 PRELIMINARY; PRT; 25 AA.

AC 016308;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

GN GLYCOPHORIN A (FRAGMENT).

OS GTPA.

OC Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95282423; PubMed=7762218;

RA Dupont B.R., Grant S.G., Oto S.H., Bigbee W.L., Jensen R.R.,

RA Langlois R.G.;

RT "Molecular characterization of glycoporphin A transcripts in human

RT erythroid cells using RT-PCR, allele-specific restriction, and

RT sequencing";

RL Vox Sang. 68:121-129(1995).

DR EMBL; S77082; AAB34408.1; -.

FT NON_TER

SQ SEQUENCE 25 AA; 2494 MW; 0DA5D769A868BEC3 CRC64;

Query Match 21.2%; Score 24; DB 4; Length 25;

Best Local Similarity 50.0%; Pred. No. 4.3e+03;

Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 13 EILHCOI 20

Db 15 EVAMHTST 22

RESULT 24

ID 09UCG1

AC 09UCG1; PRELIMINARY; PRT; 24 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE PEPTIDYL-ALPHA-HYDROXYGLYCINE ALPHA-AMIDATING LYASE (FRAGMENT).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE.

RA Husten E.J., Tausk F.A., Keutmann H.T., Elipper B.A.;

RL J. Biol. Chem. 268:9709-9717(1993).

SQ SEQUENCE 24 AA; 2830 MW; 85463403B13A147 CRC64;

Query Match 20.8%; Score 23.5; DB 4; Length 24;

Best Local Similarity 53.8%; Pred. No. 4.9e+03;

Matches 7; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 4 LADOP-ONLEIL 15

Db 3 LMOQKOGEEVL 15

RESULT 25

ID 09TUP6

AC 09TUP6; PRELIMINARY; PRT; 24 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE HEART FATY ACID BINDING PROTEIN (FRAGMENT).

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-PIETRAIN; TISSUE=BLOOD;

RA Necheleberger D., Mueller S., Mueller M.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF164968; AAD47821.1; -.

DR HSSP; P05413; 1HMT.

DR INTERPRO; IPR000463; -.

DR PFAM; PF00061; 1lipocalin; 1.

DR PROSITE; PS00214; FABP; 1.

FT NON_TER

SQ SEQUENCE 24 AA; 2782 MW; C7533989DDC2043 CRC64;

Query Match

20.8%; Score 23.5; DB 6; Length 24;

Best Local Similarity 28.6%; Pred. No. 4.9e+03;

Matches 4; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

OY 2 WELADOPONLEIL 15

Db 9 WKLVDs-KNFDDYM 21

Search completed: February 5, 2001, 10:55:24

Job time: 913 sec

QY 1 TLKYAIKTGHPRYFNQLSTG 20

```

Db      2  TLTLTIQTSXPRGMIDYNTG 21

RESULT  3
PF0256  Ig heavy chain CRD3 region (clone 2-115C) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PF0256
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PF0222; MUID:91108337
A:Accession: PF0256
A:Molecule type: DNA
A:Residues: 1-13 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match      25.9%; Score 28; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1  TLKYAKTCHPR 12
      || | : || : |
Db      1  TLYYDMLTGYSR 12

RESULT  4
PF0215  metalloendopeptidase (EC 3.4.24.-) - Streptomyces griseus (fragment)
C:Species: Streptomyces griseus
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 28-May-1993
C:Accession: PQ0215
R:Tsuuyuki, H.; Kajiura, K.; Fujita, A.; Kumazaki, T.; Ishii, S.
J. Biochem. 110, 339-344, 1991
A:Title: Purification and characterization of Streptomyces griseus metalloendopeptidases
A:Reference number: PQ0215; MUID:92121147
A:Accession: PQ0215
A:Molecule type: protein
A:Residues: 1-20 <TSU>
C:Keywords: hydrolase; metalloproteinase

Query Match      25.9%; Score 28; DB 2; Length 20;
Best Local Similarity 38.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      8  TCHPRYNQLSTG 20
      || : | | : |
Db      2  TGNQYNGQVTLG 14

RESULT  5
I64828  gene HEXA protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C:Accession: I64828
R:Boles, D.J.; Proia, R.L.
Am. J. Hum. Genet. 56, 716-724, 1995
A:Title: The molecular basis of HEXA mRNA deficiency caused by the most common Tay-Sachs
A:Reference number: I51882; MUID:95193801
A:Accession: I64828
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-25 <RES>
A:Cross-References: GB:S76982; NID:9912780; PIDN:AADL4242.1; PID:g4261942
C:Genetics:
A:Gene: GDB:HEXA
A:Cross-References: GDB:120040; OMIM:272800
A:Map position: 15q23-15q24

C:Superfamily: beta-hexosaminidase

Query Match      25.9%; Score 28; DB 2; Length 25;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      11  PRYFNQLSTG 20
      || | : || : |
Db      1  PWYLNRIISVG 10

RESULT  6
B54548  penicillin-binding protein 4 - Enterococcus hirae (fragment)
C:Species: Enterococcus hirae
C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 04-Dec-1994
C:Accession: B54548
R:Jacques, P.; el Kharroubi, A.; Van Beeumen, J.; Piras, G.; Coyette, J.; Ghuyssen, J.
FEMS Microbiol. Lett. 66, 119-123, 1991
A:Title: Mode of membrane insertion and sequence of a 32-amino acid peptide stretch o
A:Reference number: A54548; MUID:92038914
A:Accession: B54548
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <JAC>
A:Note: sequence extracted from NCBI backbone (NCBIP:63854)

Query Match      25.9%; Score 28; DB 2; Length 25;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      1  TLKYAKTCHPR 12
      || | : || : |
Db      10  TLTLTIQTSXPR 21

RESULT  7
S48156  alpha-amylase inhibitor - rye
C:Species: Secale cereale (rye)
C:Date: 07-May-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S48156
R:Garcia-Casado, G.; Sanchez-Monge, R.; Lopez-Otin, C.; Salcedo, G.
Eur. J. Biochem. 224, 525-531, 1994
A:Title: Rye inhibitors of animal alpha-amylases show different specificities, aggregat
A:Reference number: S48156; MUID:95010030
A:Accession: S48156
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-23 <GAR>
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: alpha-amylase inhibitor

Query Match      25.0%; Score 27; DB 2; Length 23;
Best Local Similarity 80.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      9  GHPRY 13
      || | |
Db      11  GHPMY 15

RESULT  8
PH1728  Ig heavy chain V region (clone GCC-8) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1728
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993

```



```

Query Match      23.1%; Score 25; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. NO. 8.6e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      8 TGHPRYFN 15
      ||| : |
Db       9 TGHEEFSN 16

```

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Accession: U00001
A:Residues: 1-12 <RES>
A:Cross-references: GB:S76984; NID:G912781; PIDN:AAD14243.1; PID:94261943
C:Genetics:
A:Gene: GDB:HEXA
A:Cross-references: GDB:120040; OMIM:272800
A:Map position: 15q23-15q24

C:Superfamily: beta-hexosaminidase

Query Match 22.2%; Score 24; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 8.7e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 11 PRYFNQLS 18

Db 1 PWYLNRI 8

RESULT 14

PH1464

T-cell receptor beta chain (clone A3/63) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995

C:Accession: PH1464

R:Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Kojima, Y.; Exp. Med. 177, 811-820, 1993

A:Title: T cell receptor selection by and recognition of two class I major histocompatibility complex peptides

A:Reference number: PH1430; MUID:93171821

A:Accession: PH1464

A:Molecule type: mRNA

A:Residues: 1-12 <CAS>

A:Experimental source: cytolytic T-lymphocyte

C:Superfamily: immunoglobulin homology

C:Keywords: receptor; T-cell

Query Match 22.2%; Score 24; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 8.7e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 AIKTGHPRYF 14

Db 2 ASSTGNTLYF 11

RESULT 15

H56046

urinary tract stone matrix protein 10, 42K - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 12-Apr-1995

C:Accession: H56046

R:Binette, J.P.; Binette, M.B.; Gawinowicz, M.A.; Kendrick, N.

A:Title: Isolation, characterization and sequence of stone proteins.

A:Description: Isolation, characterization and sequence of stone proteins.

A:Reference number: A56046

A:Accession: H56046

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-13 <BIN>

Query Match 22.2%; Score 24; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 9.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 13 YFNQLS 18

Db 4 YFNDLA 9

RESULT 16

B41868

hypothetical protein (traE1 3' region) - Enterococcus faecalis plasmid pAD1

C:Species: Enterococcus faecalis

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999

C:Accession: B41868; B37391

R:Pontius, L.T.; Clewell, D.B.

J. Bacteriol. 174, 3152-3160, 1992

A:Title: Conjugative transfer of Enterococcus faecalis plasmid pAD1: nucleotide sequence

A:Reference number: A41868; MUID:92250408

A:Contents: plasmid pAD1

A:Accession: B41868

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-15 <PON>

A>Note: sequence extracted from NCBI backbone (NCBIN:99901, NCBIP:99906)

R:Clewell, D.B.; Pontius, L.T.; An, F.Y.; Ike, Y.; Suzuki, A.; Nakayama, J.

Plasmid 24, 156-161, 1990

A:Title: Nucleotide sequence of the sex pheromone inhibitor (iAD1) determinant of Ent

A:Reference number: A37391; MUID:91261999

A:Accession: B37391

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-15 <CLE>

A:Cross-references: GB:M62888; NID:g141853; PIDN:AAA98040.1; PID:g141855

C:Genetics:

A:Genome: plasmid

Query Match 22.2%; Score 24; DB 2; Length 15;

Best Local Similarity 36.4%; Pred. No. 1.1e+03;

Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 YAIKTGHPRYF 14

Db 2 YTVHVYIPRF 12

RESULT 17

A61612

allotostatin - tobacco hornworm

C:Species: Manduca sexta (tobacco hornworm)

C>Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 11-Jul-1997

C:Accession: A61612

R:Kramer, S.J.; Toschi, A.; Miller, C.A.; Kataoka, H.; Quistad, G.B.; Li, J.P.; Carne

Proc. Natl. Acad. Sci. U.S.A. 88, 9458-9462, 1991

A:Title: Identification of an allotostatin from the tobacco hornworm Manduca sexta.

A:Reference number: A61612; MUID:92052112

A:Accession: A61612

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-15 <KRA>

C:Keywords: neuropeptide; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 22.2%; Score 24; DB 2; Length 15;

Best Local Similarity 66.7%; Pred. No. 1.1e+03;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 13 YFNQLS 18

Db 8 YFNPI 13

RESULT 18

PH1304

Ig heavy chain DJ region (clone C439-111) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1304

R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A:Title: Predominance of fetal type DJH joining in young children with B precursor ly

A:Reference number: PH1302; MUID:93094761

A:Accession: PH1304

A:Molecule type: DNA

A:Residues: 1-19 <WAS>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 22.2%; Score 24; DB 2; Length 19;
Best Local Similarity 42.9%; Pred. No. 1.4e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 TGHPRYF 14
| | | |
Db 1 SGYPYY 7

RESULT 19
S58382
hypothetical protein 1 - human
C:Species: Homo sapiens (man)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 08-Oct-1999
C:Accession: S58382
R:Dirks, R.P.H.; Onnekink, C.; Jansen, H.J.; de Jong, A.; Bloemers, H.P.J.
Nucleic Acids Res. 23, 2815-2822, 1995
A:Title: A novel human c-sis mRNA species is transcribed from a promoter in c-sis intron
A:Reference number: S58382; MUID:95388493
A:Accession: S58382
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-20 <DIR>
A:Cross-references: EMBL:X83705; NID:g951023; PIDN:CAA58678.1; PID:g951024

Query Match 22.2%; Score 24; DB 2; Length 20;
Best Local Similarity 46.2%; Pred. No. 1.5e+03;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 LKVAIKTGHPRYF 14
| | | | | |
Db 5 LSWASGTFFPRSF 17

RESULT 20
S47198
T-cell receptor J-alpha wIII.2 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 05-Nov-1999
C:Accession: S47198
R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A:Reference number: S40133
A:Accession: S47198
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-21 <PLA>
A:Cross-references: EMBL:X71041; NID:g507005; PIDN:CAA50358.1; PID:g510655
C:Keywords: T-cell receptor

Query Match 22.2%; Score 24; DB 2; Length 21;
Best Local Similarity 55.6%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLKYAIKGTG 9
| | | |
Db 6 TYKYIFGTG 14

RESULT 21
S47212
T-cell receptor J-alpha wIII.3 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 05-Nov-1999
C:Accession: S47212
R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A:Reference number: S40133
A:Accession: S47212
A>Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-21 <PLA>
A:Cross-references: EMBL:X71043; NID:g506512; PIDN:CAA50360.1; PID:g510316
C:Keywords: T-cell receptor

Query Match 22.2%; Score 24; DB 2; Length 21;
Best Local Similarity 55.6%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLKYAIKGTG 9
| | | |
Db 6 TYKYIFGTG 14

RESULT 22
S04228
N4(beta-N-acetylglucosaminyl)-L-asparaginase (EC 3.5.1.26) 20K chain - rat (fragment)
N:Alternate names: glycosylasparaginase
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 28-Apr-1993
C:Accession: S04228
R:Tollersrud, O.K.; Aronson Jr., N.N.
Biochem. J. 260, 101-108, 1989
A:Title: Purification and characterization of rat liver glycosylasparaginase.
A:Reference number: S04228; MUID:89374025
A:Accession: S04228
A:Molecule type: protein
A:Residues: 1-22 <TOL>
C:Keywords: hydrolase

Query Match 22.2%; Score 24; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 KTGH 10
| | | |
Db 9 KTGH 12

RESULT 23
PH1907
T-cell receptor alpha chain (clone A21) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 17-Mar-1999
C:Accession: PH1907
R:Sensi, M.; Salvi, S.; Castelli, C.; Maccalli, C.; Mazzocchi, A.; Mortarini, R.; Nic
J. Exp. Med. 178, 1231-1246, 1993
A:Title: T cell receptor (TCR) structure of autologous melanoma-reactive cytotoxic T
HLA-A2-restricted and melanocyte-lineage-specific CTL clone.
A:Reference number: PH1907; MUID:93389388
A:Accession: PH1907
A:Molecule type: mRNA
A:Residues: 1-25 <SEN>
A:Cross-references: EMBL:X74392
A:Experimental source: lymphocyte
C:Keywords: receptor

Query Match 22.2%; Score 24; DB 2; Length 25;
Best Local Similarity 55.6%; Pred. No. 1.9e+03;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLKYAIKGTG 9
| | | |
Db 7 TYKYIFGTG 15

RESULT 24
S42587
celf protein - Escherichia coli (fragment)
C:Species: Escherichia coli
C:Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995

Search completed: February 5, 2001, 10:49:30
Job time: 743 sec

C:Accession: S42587
R:Guzzo, A.; DuBow, M.S.
Mol. Gen. Genet. 242, 455-460, 1994
A:Title: A luxAB transcriptional fusion to the cryptic celf gene of Escherichia coli
A:Reference number: S42587; MUID:94166755

A:Accession: S42587
A:Molecule type: DNA
A:Residues: 1-11 <GUZ>
C:Genetics:
A:Gene: celf

Query Match 21.3%; Score 23; DB 2; Length 11;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 TCHPR 12
| | | |
DB 6 TPHPR 10

RESULT 25

S47355
T-cell antigen receptor VJ junction beta chain - human
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S47365; S47375; S47396; S47397; S47398; S47355
R:Lehner, P.J.

submitted to the EMBL Data Library, August 1994

A:Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T
A:Reference number: S47355
A:Accession: S47365
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-13 <LEH>
A:Cross-references: EMBL:Z35690; NID:527471; PIDN:CAA84759.1; PID:9527472; EMBL:Z35679;

A:Accession: S47375
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-13 <LE2>
A:Cross-references: EMBL:Z35700; NID:527493; PIDN:CAA84769.1; PID:9527494

A:Accession: S47379
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-13 <LE3>
A:Cross-references: EMBL:Z35708; NID:527509; PIDN:CAA84777.1; PID:9527510

A:Accession: S47396
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-13 <LE4>
A:Cross-references: EMBL:Z35674; NID:527527; PIDN:CAA84743.1; PID:9527528

A:Accession: S47397
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-13 <LE5>
A:Cross-references: EMBL:Z35675; NID:527529; PIDN:CAA84744.1; PID:9527530

A:Accession: S47398
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-13 <LE6>
A:Cross-references: EMBL:Z35676; NID:527531; PIDN:CAA84745.1; PID:9527532
C:Keywords: T-cell receptor

Query Match 21.3%; Score 23; DB 2; Length 13;
Best Local Similarity 30.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 5 AKTGHPRYF 14
| | | | |
DB 4 SIRSYEQYF 13

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:56:00 ; Search time 32.57 Seconds
(without alignments)
19.831 Million cell updates/sec

Title: US-08-981-824-4

Perfect score: 108
Sequence: 1 TLKVAIKTGHPRYFNQLSTG 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 1436

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	25.0	17	1	RANR_RANRU
2	25	23.1	19	1	DHAB_COMTE
3	24	22.2	15	1	ALLS_MANSE
4	23	21.3	9	1	OXYF_SCYCA
5	23	21.3	16	1	FORI_MYRGU
6	23	21.3	19	1	PPBH_PSEAE
7	23	21.3	21	1	CSPS_STRTR
8	23	21.3	24	1	PGQ_XENXA
9	22	20.4	15	1	CBPB_PROAT
10	22	20.4	19	1	FRHG_METBA
11	22	20.4	19	1	OXLX_OPHHA
12	22	20.4	25	1	PKI2_SOLTU
13	21	19.4	4	1	DCML_PSECH
14	21	19.4	15	1	AH2_PRUSE
15	21	19.4	18	1	CYCH_MOUSE
16	21	19.4	18	1	MLB_SCYCA
17	21	19.4	19	1	HBB2_URCHA
18	21	19.4	21	1	DCMS_PSECA
19	20	18.5	9	1	NSK1_SARBU
20	20	18.5	10	1	LSK2_LEUMA
21	20	18.5	11	1	LSK1_LEUMA
22	20	18.5	11	1	LSKP_PERAM
23	20	18.5	12	1	LOSK_LOOMI
24	20	18.5	14	1	NSK2_SANBU
25	20	18.5	15	1	UCO8_MAIZE
26	20	18.5	16	1	MLB_SOUNC
27	20	18.5	17	1	APID_BOMPA
28	20	18.5	17	1	PROP_SATY
29	20	18.5	20	1	FIBR_PACLE
30	20	18.5	20	1	MIF_PIG
31	20	18.5	20	1	OMPL_ACTAC
32	20	18.5	22	1	ATP6_COTJA
33	20	18.5	22	1	KORR_MERTM

34	20	18.5	22	1	SETB_SATY	P33077 salmoneila
35	20	18.5	24	1	COXC_THOIB	P80973 thunnus obe
36	19.5	18.1	22	1	RLI8_HALME	P50562 halobacteri
37	19.5	18.1	23	1	RLI8_HALSA	P50562 halobacteri
38	19.5	18.1	23	1	RLI8_HALVO	P50563 halobacteri
39	19	17.6	9	1	LITR_PHTRO	P08946 phyllomedus
40	19	17.6	10	1	ANG1_BOTJA	Q10581 bothrops ja
41	19	17.6	11	1	RANC_RANPI	P08951 rana pipien
42	19	17.6	11	1	RR2_CONAM	P42341 conopholis
43	19	17.6	11	1	TKN3_PSEGU	P42988 pseudophryn
44	19	17.6	14	1	LPF2_ECOLI	P06985 escherichia
45	19	17.6	15	1	FGF1_CANFA	P18651 canis famli

ALIGNMENTS

RESULT 1	RANR_RANRU	STANDARD;	PRT;	17 AA.
ID	RANR_RANRU			
AC	P08952;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	01-NOV-1988 (Rel. 09, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	RANA_RUGOSA (Frog).			
OS	Rana rugosa (Frog).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.			
RN	[1]			
RP	SEQUENCE			
RX	MEDLINE-84131098; PubMed-6141890;			
RU	Nakajima T.;			
RU	Unpublished results, cited by:			
RU	Espartero V., Espartero G.F., Mazzanti G., Endean R.;			
RU	Comp. Biochem. Physiol. 77C:99-108(1984).			
CC	-1- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROKININ B/RANATENSIN			
CC	FAMILY.			
CC	INTERPRO: IPR000874; -			
DR	PFAM: PF02044; Bombesin; 1.			
DR	PROSITE: PS00257; BOMBESIN; 1.			
KW	Bombesin family; Amidation.			
FT	MOD_RES 17			
FT	SEQUENCE 17 AA; 2053 MW; 3A876B35A581863E CRC64;			
QY	12 RYFNQLSTG 20			
Db	6 RRYNQWATG 14			
Query Match	25.0%; Score 27; DB 1; Length 17;			
Best Local Similarity	55.6%; Pred. No. 1.8e+02; Mismatches 2; Indels 0; Gaps 0;			
Matches	5; Conservative			
RESULT 2	DHAB_COMTE	STANDARD;	PRT;	19 AA.
ID	DHAB_COMTE			
AC	P80704;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	ALDEHYDE DEHYDROGENASE, BETA CHAIN (EC 1.2.99.3) (ALDH) (FRAGMENT).			
OS	Comamonas testosteroni (Pseudomonas testosteroni).			
OC	Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.			
RN	[1]			
RP	SEQUENCE			
RC	STRAIN-ATCC 15667;			
RA	Luykx D.M.A.M., Kim S.W., de Vries S., Duine J.A.;			
RU	Submitted (JUL-1996) to the SWISS-PROT data bank.			
CC	-1- CATALYTIC ACTIVITY: AN ALDEHYDE + ACCEPTOR + H(2)O -> AN ACID + REDUCED ACCEPTOR.			
CC	-1- COFACTOR: MOLYBDENUM.			
CC	-1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA AND A GAMMA.			

CC CHAIN.
 KM Oxioreductase; Molybdenum.
 FT NON_TER 19
 SQ SEQUENCE 19 AA; 2096 MW; 89BD67DAD05A212E CRC64;

Query Match 23.1%; Score 25; DB 1; Length 19;
 Best Local Similarity 33.3%; Pred. No. 4.2e+02;
 Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 4 YAIKTGHPRENQLS 18
 DB 2 YAFSTPRTLDEV 16

RESULT 3
 ALLS MANSE STANDARD; PRT; 15 AA.
 AC P42559;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ALLATOSTATIN (MAS-AS).
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-HEAD;
 RX MEDLINE-92052112; PubMed-1946359;
 RA Kramer S.J., Toschl A., Miller C.A., Katooka H., Qulstad G.B.,
 RA Li J.P., Carney R.L., Schooley D.A.;
 RT "Identification of an allatostatin from the tobacco hornworm Manduca
 sexta";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9458-9462(1991).
 CC -1- FUNCTION: STRONGX INHIBITS JUVENILE HORMONE BIOSYNTHESIS IN VITRO
 BY THE CORPORA ALLATA FROM FIFTH-STADIUM LARVAE AND ADULT FEMALES.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 CC Neuropeptide.
 KW MOD.RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD.RES 1 1
 SQ SEQUENCE 15 AA; 1908 MW; 1605B77CDEBC838E CRC64;

Query Match 22.2%; Score 24; DB 1; Length 15;
 Best Local Similarity 66.7%; Pred. No. 4.7e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 13 YFNQLS 18
 DB 8 YFNPIIS 13

RESULT 4
 OXYF-SCYCA STANDARD; PRT; 9 AA.
 ID OXYF-SCYCA
 AC P42997;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE PHASVANOCIN
 OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
 OC Eukaryota; Metazoa; Chordata; Craniata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphi; Galeoidea; Carcharhiniformes;
 OC Scyllorhinidae; Scyllorhinus.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-PITUITARY;
 RX MEDLINE-95062247; PubMed-7972045;
 RA Chauvet J., Rouille Y., Chauveau C., Chauvet M.-T., Acher R.;
 RA "Special evolution of neurohypophyseal hormones in cartilaginous
 fishes: asatocin and phasvatocin, two oxytocin-like peptides
 isolated from the spotted dogfish (Scyllorhinus caniculus).";

RL Proc. Natl. Acad. Sci. U.S.A. 91:11266-11270(1994).
 CC -1- FUNCTION: DISPLAYS OXYTOCIC ACTIVITY ON RAT UTERS.
 CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
 DR INTERPRO: IPR000981;
 DR PFAM: PF00220; hormone4; 1.
 DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
 KM Hormone; Amidation.
 FT DISULFID 1 6
 FT MOD.RES 9 9 AMIDATION.
 SQ SEQUENCE 9 AA; 1016 MW; 17EDD76EB44449DB CRC64;

Query Match 21.3%; Score 23; DB 1; Length 9;
 Best Local Similarity 50.0%; Pred. No. 8.9e+04;
 Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 13 YFNQSTG 20
 DB 2 YFNCPVG 9

RESULT 5
 FORL MYRGU STANDARD; PRT; 16 AA.
 ID FORL MYRGU
 AC P81438;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE FORMACIN 1.
 OS Myrmecia gulosa (Red bulldog ant).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
 OC Formicidae; Formicidae; Myrmecia.
 RN [1]
 RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
 RC TISSUE-HEMOLYMPH;
 RX MEDLINE-98165787; PubMed-9497332;
 RA MacIntosh J.A., Veal D.A., Beattie A.J., Gooley A.A.;
 RT "Isolation from an ant Myrmecia gulosa of two inducible
 O-glycosylated proline-rich antibacterial peptides";
 RL J. Biol. Chem. 273:6139-6143(1998).
 CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE. HAS ACTIVITY AGAINST E. COLI
 BUT NONE AGAINST OTHER GRAM-NEGATIVE BACTERIA AND GRAM-POSITIVE
 BACTERIA.
 CC -1- INDUCTION: UPON BACTERIAL CHALLENGE.
 CC -1- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
 CC GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
 CC -1- SIMILARITY: TO DROSOPHILA DROSOCIN.
 CC Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
 KW CARBOHYD 11 11 O-LINKED (GALNAC...)
 FT CARBOHYD 11 11
 SQ SEQUENCE 16 AA; 1794 MW; 80CEA3AABBC2E0AE CRC64;

Query Match 21.3%; Score 23; DB 1; Length 16;
 Best Local Similarity 80.0%; Pred. No. 7.3e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 8 TGHPR 12
 DB 11 TPHPR 15

RESULT 6
 PPBH_PSEAE STANDARD; PRT; 19 AA.
 ID PPBH_PSEAE
 AC P35483;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ALKALINE PHOSPHATASE H (EC 3.1.3.1) (H-AP) (FRAGMENT).
 OS Pseudomona aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.

RN [1]
 RP SEQUENCE.
 AC STRAIN-H103;
 RX MEDLINE-93202452; PubMed-8454193;
 RA Tan A.S.P., Morobec E.A.;
 RT "Isolation and characterization of two immunologically distinct
 alkaline phosphatases from *Pseudomonas aeruginosa*."
 RL FEMS Microbiol. Lett. 106:281-286(1993).
 CC -1- FUNCTION: HAS ONLY A PHOSPHOMONESTERASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE (AT A HIGH PH OPTIMUM)
 CC -1- COFACTOR: METALLOENZYME CONTAINING TWO ZINC ATOMS AND A MAGNESIUM
 CC ION.
 CC -1- SUBCELLULAR LOCATION: SECRETED AND PERIPLASMIC.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED CONSTITUTIVELY.
 CC -1- SIMILARITY: BELONGS TO THE ALKALINE PHOSPHATASE FAMILY.
 DR INTERPRO: IPR001952; .
 DR PROSITE: PS00123; ALKALINE_PHOSPHATASE: PARTIAL.
 KM Hydrolyase; Zinc; Magnesium; Periplasmic.
 FT NON_TER 19
 SQ SEQUENCE 19 AA; 2131 MW; C51B09D7DB22E799 CRC64;

Query Match 21.3%; Score 23; DB 1; Length 19;
 Best Local Similarity 40.0%; Pred. No. 8.8e+02;
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 11 PYYENLSTG 20
 Db 5 PSLFNRQAG 14

RESULT 7
 ID CSFS_STRTR STANDARD; PRT; 21 AA.
 AC P81632;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE COLD SHOCK PROTEIN CSFPT (FRAGMENT).
 OS Streptococcus thermophilus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 RN [1]
 RP SEQUENCE.
 RC STRAIN-P818;
 RX MEDLINE-99456673; PubMed-10525839;
 RA Perrin C., Guilmont C., Bracquart P., Galliard J.L.;
 RT "Expression of a new cold shock protein of 21.5 kDa and of the major
 cold shock protein by *Streptococcus thermophilus* after cold shock."
 RL Curr. Microbiol. 39:342-347(1999).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- INDUCTION: BY COLD SHOCK.
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
 DR INTERPRO: IPR002059; .
 DR PRAM: PF00313; CSD; 1.
 DR PROSITE: PS00352; COLD_SHOCK: PARTIAL.
 KM Transcription regulation; DNA-binding; Activator.
 FT DOMAIN 1 >21
 FT NON_TER 21
 SQ SEQUENCE 21 AA; 2376 MW; 56A4A2800F3345EA CRC64;

Query Match 21.3%; Score 23; DB 1; Length 21;
 Best Local Similarity 44.4%; Pred. No. 9.7e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 KTGHPKRFN 15
 Db 1 KNGTKWFN 9

RESULT 8

PGQ_XENLA
 ID PGQ_XENLA STANDARD; PRT; 24 AA.
 AC P39080;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE ANTIMICROBIAL PEPTIDE PGQ.
 OS *Xenopus laevis* (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; *Xenopus*.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-STOMACH;
 RX MEDLINE-92011794; PubMed-1717472;
 RA Moore K.S., Bevins C.L., Brasseur M.M., Tomassini N., Turner K.,
 RA Eck H., Zasloff M.;
 RT "Antimicrobial peptides in the stomach of *Xenopus laevis*."
 RL J. Biol. Chem. 266:19851-19857(1991).
 CC -1- FUNCTION: ANTIMICROBIAL PEPTIDE.
 CC -1- TISSUE SPECIFICITY: IS SYNTHESIZED IN THE STOMACH AND STORED
 CC IN A NOVEL GRANULAR MULTINUCLEATED CELL IN THE GASTRIC MUCOSA.
 CC IT IS STORED AS ACTIVE, PROCESSED PEPTIDES IN LARGE GRANULES
 CC WITHIN THE GRANULAR GLAND SECRETIONS OF THE SKIN.
 CC -1- SIMILARITY: BELONGS TO THE MAGAINTIN FAMILY OF ANTIMICROBIAL
 CC PEPTIDES.
 DR PIR: A41037; A41037.
 KM Antibiotic; Amphibian skin.
 FT Antibiotic; Amphibian skin.
 SQ SEQUENCE 24 AA; 2457 MW; 76GA87CB7CF22B9C CRC64;

Query Match 21.3%; Score 23; DB 1; Length 24;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 13 YFNOLSTG 20
 Db 9 YLKNLGTG 16

RESULT 9
 ID CBPR_PROAT STANDARD; PRT; 15 AA.
 AC P19628;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE CARBOXYPEPTIDASE B (EC 3.4.17.2) (FRAGMENT).
 OS *Protophytes aethiops* (Marbled lungfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Dipnoi; Lepidosireniformes; Protopteridae; Protopterus.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-PANCREAS;
 RX MEDLINE-73025047; PubMed-5079891;
 RA Reeck G.R., Neutath H.;
 RT "Isolation and characterization of pancreatic procarboxypeptidase B
 and carboxypeptidase B of the African lungfish."
 RL Biochemistry 11:3947-3955(1972).
 CC -1- CATALYTIC ACTIVITY: PEPTIDYL-L-LYSINE/ARGININE + H(2)O -> PEPTIDE +
 CC L-LYSINE/ARGININE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M14; ALSO KNOWN AS THE
 CC ZINC CARBOXYPEPTIDASE FAMILY.
 CC PIR: A26212; A26212.
 DR MEROPS: M14.003; .
 DR INTERPRO: IPR000834; .
 DR PROSITE: PS00132; CARBOXYPEPT_2N_1; PARTIAL.
 DR PROSITE: PS00133; CARBOXYPEPT_2N_2; PARTIAL.
 KM Hydrolyase; Carboxypeptidase; Zinc; Zymogen.
 FT PROPEP 1 >15
 FT NON_TER 15
 SQ SEQUENCE 15 AA; 1749 MW; 124C910D937BED65 CRC64;

Query Match 20.4%; Score 22; DB 1; Length 15;

Best Local Similarity 80.0%; Pred. No. 9.8e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 11 PRYFN 15

DB 4 PRSFN 8

RESULT 10

FRHG_METBA

AC FRHG_METBA

DR 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DE COENZYME F420 HYDROGENASE GAMMA SUBUNIT (EC 1.12.99.1) (8-HYDROXY-5-DEAZAFLAVIN-REDUCING HYDROGENASE GAMMA SUBUNIT) (FRAGMENT).

GN FRHG.

OS Methanosarcina barkeri.

OC Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;

CC Methanosarcina.

RN [1]

RP SEQUENCE.

RC STRAIN-FUSARO / DSM 804;

RX MEDLINE-96085134; PubMed-8521835;

RA Michel R., Massanz C., Kostka S., Richter M., Fiebig K.;

RT "Biochemical characterization of the 8-hydroxy-5-deazaflavin-reactive hydrogenase from Methanosarcina barkeri Fusaro.";

RL Eur. J. Biochem. 233:727-735(1995).

CC -1- FUNCTION: REDUCES THE PHYSIOLOGICAL LOW-POTENTIAL TWO-ELECTRON ACCEPTOR COENZYME F420, AND THE ARTIFICIAL ONE-ELECTRON ACCEPTOR METHYLVIOLIGEN.

CC -1- CATALYTIC ACTIVITY: H(2) + COENZYME F420 - REDUCED COENZYME F420.

CC -1- COFACTOR: FHN CONTAINS NICKEL, IRON-SULFUR, AND FAD COFACTORS.

CC -1- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS AND A GAMMA CHAIN.

CC -1- SIMILARITY: TO THE SMALL SUBUNITS OF OTHER NI-CONTAINING HYDROGENASES.

CC Oxidoreductase; Iron-sulfur; Electron transport.

KW NON-TER 19

FT SEQUENCE 19 AA; 2012 MW; C443B09B9E7B9D58 CRC64;

SO

Query Match 20.4%; Score 22; DB 1; Length 19;

Best Local Similarity 80.0%; Pred. No. 1.3e+03;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 IKTGH 10

DB 4 IKIGH 8

RESULT 11

OXA_OPNHA

AC OXA_OPNHA

DR 15-DEC-1998 (Rel. 37, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DE 30-MAY-2000 (Rel. 39, Last annotation update)

DE L-AMINO ACID OXIDASE (EC 1.4.3.2) (LAO) (LAO) (FRAGMENT).

OS Ophiophagus hannah (King cobra) (Naja hannah).

OC Ophiophaga; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;

CC Elapidae; Elapinae; Ophiophagus.

RN [1]

RP SEQUENCE.

RC TISSUE=VENOM;

RX MEDLINE-94361525; PubMed-8080286;

RA Ponnudurai G., Chung M.C.M., Tan N.-H.;

RT "Purification and properties of the L-amino acid oxidase from Malayan pit viper (Calloselasma rhodostoma) venom.";

RL

Query Match 20.4%; Score 22; DB 1; Length 25;

Best Local Similarity 50.0%; Pred. No. 1.7e+03;

Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 9 GHPYFNQ 16

DB 11 GHPYRIGQ 18

RESULT 13

RL Arch. Biochem. Biophys. 313:373-378(1994).

RN [2]

RP SEQUENCE OF 1-15.

RC TISSUE=VENOM;

RX MEDLINE-97449790; PubMed-9304806;

RA Ahn M.Y., Lee B.M., Kim Y.S.;

RT "Characterization and cytotoxicity of L-amino acid oxidase from the venom of king cobra (Ophiophagus hannah).";

RL Int. J. Biochem. Cell Biol. 29:911-919(1997).

CC -1- FUNCTION: HAS CYTOTOXIC ACTIVITY (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: A L-AMINO ACID + H(2)O + O(2) - A 2-OXO ACID + NH(3) + H(2)O(2).

CC -1- COFACTOR: FAD.

CC -1- SUBUNIT: HOMODIMER (PROBABLE).

CC -1- PTM: GLYCOSYLATED.

CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOXAMINE OXIDASE FAMILY.

CC STRONG, TO MOUSE FIG-1.

CC Oxidoreductase; Flavoprotein; FAD; Glycoprotein; Venom.

CC CONFLICT 1

FT NON-TER 19

FT SEQUENCE 19 AA; 2298 MW; DD91A5B414F1427 CRC64;

SO

Query Match 20.4%; Score 22; DB 1; Length 19;

Best Local Similarity 50.0%; Pred. No. 1.3e+03; Mismatches 3; Indels 0; Gaps 0;

OY 11 PRYFNQLS 18

DB 12 PRYFNHMA 19

RESULT 12

ID PK12_SOLTU

AC PK12_SOLTU

DR 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DE KUNITZ-TYPE INHIBITOR-2 (PK1-2) (FRAGMENT).

OS Solanum tuberosum (Potato).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

CC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; easterids I;

CC Solanales; Solanaceae; Solanum.

RN [1]

RP SEQUENCE.

RC STRAIN=CV. RUSSET BURBANK; TISSUE=TUBER;

RA Walsh T.A., Twichell W.P.;

RT "Two Kunitz-type proteinase inhibitors from potato tubers.";

RL Plant Physiol. 97:15-18(1991).

CC -1- FUNCTION: POTENT INHIBITOR OF SUBTILISIN. MODERATE INHIBITOR OF TRYPSIN AND CHYMOTRYPSIN.

CC -1- TISSUE SPECIFICITY: CORTEX OF POTATO TUBER.

CC -1- SIMILARITY: TO SOYBEAN TRYPSIN INHIBITOR (KUNITZ) FAMILY OF PROTEASE INHIBITOR.

CC INTERPRO: IPR002160; -

DR PFAM: PF00197; Kunitz_legume; 1.

DR PROSITE: PS00283; SOYBEAN_KUNITZ; 1.

DR Kunitz-type proteinase inhibitor.

FT NON-TER 25

FT SEQUENCE 25 AA; 2920 MW; 1F5607405921DFDC CRC64;

SO


```

DCML_PSECH
ID DCML_PSECH STANDARD: PRT: 4 AA.
AC P19916;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE CARBON MONOXIDE DEHYDROGENASE LARGE CHAIN (EC 1.2.99.2) (FRAGMENT)
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria.
RN [1]
RP MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
  carboxydotrophic bacteria."
RL Arch. Microbiol. 152:335-341(1989).
CC -1- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR - CO(2) + REDUCED
  ACCEPTOR
CC -1- COFACTOR: MOLYBDENUM.
CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
  SMALL.
DR PIR: P10140; P10140.
KM Oxidoreductase: Molybdenum.
FT NON_TER 4
SQ SEQUENCE 4 AA: 441 MW: 7761876F0000000 CRC64:

Query Match 19.4%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 GHP 11
Db 2 GHP 4

RESULT 14
AH2_PRUSE
ID AH2_PRUSE STANDARD: PRT: 15 AA.
AC P29260;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE AMGDALIN BETA-GLUCOSIDASE I (EC 3.2.1.17) (AMGDALIN HYDROLASE
  ISOZYME I') (AH I') (FRAGMENT).
OS Prunus serotina (Black cherry).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
  Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
  Rosales; Rosaceae; Prunus.
RN [1]
RP SEQUENCE.
RC TISSUE=SEED;
RA Li C.P., Swain E., Poulton J.E.;
RT "Prunus serotina amygdalin hydrolase and prunasin hydrolase."
RL Plant Physiol. 100:282-290(1992)
CC -1- CATALYTIC ACTIVITY: (R)-AMGDALIN + H(2)O - (R)-PRUNASIN +
  D-GLUCOSE.
CC -1- SUBUNIT: MONOMER.
CC -1- DEVELOPMENTAL STAGE: ABSENT FROM MATURING BLACK CHERRY FRUITS
  UNTIL 6 WEEKS AFTER FLOWERING. THEN, CONCOMITANT WITH COTYLEDON
  DEVELOPMENT, THE LEVEL OF ENZYME INCREASES WITH SPECIFICITY FOR
  EMBRYONAL TISSUES.
CC -1- PFM: GLYCOSYLATED.
KM Glycosidase: Hydrolyzed; Glycoprotein; Multigene family.
FT NON_TER 15
SQ SEQUENCE 15 AA: 1650 MW: F7CC4FA321E12EC4 CRC64:

Query Match 19.4%; Score 21; DB 1; Length 15;
Best Local Similarity 45.5%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 7 KTGHRFYNOL 17

```

```

Db 2 KTDPPHFASL 12

RESULT 15
CYCH_MOUSE
ID CYCH_MOUSE STANDARD: PRT: 18 AA.
AC O61458;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYCLIN H (FRAGMENT).
GN CCNH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=TESTIS;
RA Hall F.L., Wu L.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MEMBER OF CAK WHICH ACTIVATES CYCLIN-ASSOCIATED
  CDC2/CDK2/CDK4 BY THREONINE PHOSPHORYLATION. ITS EXPRESSION AND
  ACTIVITY ARE CONSTANT THROUGHOUT THE CELL CYCLE. CAK IS TIGHTLY
  ASSOCIATED WITH A MULTIPROTEIN COMPLEX TPPIH, WHICH PLAYS A DUAL
  ROLE IN TRANSCRIPTION AND DNA REPAIR (BY SIMILARITY).
CC -1- SUBUNIT: MAMMALIAN CAK CONTAINS THREE COMPONENTS: CDK7, CYCLIN H,
  AND AN ASSEMBLY FACTOR CALLED MAT1.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN C SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
  or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: X82441; CA57822.1;
DR HSEB: P51946; IKXU.
DR INTERPRO: IPR000553;
DR PROSITE: PS00292; CYCLINS, PARTIAL.
KM Cyclin; Cell cycle; Cell division; Nuclear protein;
  Transcription regulation.
FT NON_TER 1
SQ SEQUENCE 18 AA: 2105 MW: 92964DCFB68B98C7 CRC64:

Query Match 19.4%; Score 21; DB 1; Length 18;
Best Local Similarity 36.4%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 LKVAIKTGHR 12
Db 1 LNSVMEXHPR 11

RESULT 16
MLB_SCYCA
ID MLB_SCYCA STANDARD: PRT: 18 AA.
AC P01206;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE MELANOTROPIN BETA (BETA-MSH).
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Chondrichthyes;
  Elasmobranchii; Galeomorphi; Galeidae; Carcharhiniformes;
  Scyliorhinidae; Scyliorhinus.
RN [1]
RP SEQUENCE.

```

RX MEDLINE-75113445; PubMed-4452470;
 RA Love R.M., Pickering B.T.;
 RT "A beta-MSH in the pituitary gland of the spotted dogfish
 (Scyliorhinus canicula): isolation and structure.";
 RL Gen. Comp. Endocrinol. 24:398-404(1974).
 DR PIR: A01470; MTEBEC.
 KW Hormone.
 SQ SEQUENCE 18 AA; 2195 MW; 353C93A30C314DF4 CRC64;

Query Match 19.4%; Score 21; DB 1; Length 18;
 Best Local Similarity 57.1%; Pred. No. 1.7e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 KTGHPRY 13
 DB 6 KMGHFRM 12

RESULT 17
 HBB2_UROHA STANDARD; PRT; 19 AA.
 AC P18992;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-MAY-1991 (Rel. 18, Last annotation update)
 DE HEMOGLOBIN BETA-2 CHAIN (FRAGMENT).
 OS Uromastix hardwickii (Indian spiny-tailed lizard).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Leioleptinae;
 OC Uromastix.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-84029159; PubMed-6628672;
 RA Nagy S., Zaidi Z.H., von Bahr-Lindstrom H., Carlquist M.,
 RA Joernvall H.;
 RT "Characterization of hemoglobin from the lizard Uromastix
 hardwickii.";
 RT FEBS Lett. 162:290-295(1983).
 DR PIR: A05305; A05305.
 DR INTERPRO: IPR000971; -
 DR PROSITE: PS01033; GLOBIN; PARTIAL.
 KW Hem: Oxygen transport; Respiratory protein; Erythrocyte.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;

Query Match 19.4%; Score 21; DB 1; Length 19;
 Best Local Similarity 60.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 8 TGHPR 12
 DB 15 TGNPK 19

RESULT 18
 DCMS_PSECA STANDARD; PRT; 21 AA.
 ID DCMS_PSECA
 AC P19921;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE CARBON MONOXIDE OXYGENASE [CYTOCHROME B-561] SMALL CHAIN (EC 1.2.2.4)
 DE (FRAGMENT).
 OS Pseudomonas carboxydovorans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bradyrhizobium group; Oligotropa.
 RN [1]
 RP SEQUENCE.
 RX STRAIN-OM5;
 RX MEDLINE-90055678; PubMed-2818128;

RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
 RT "Homology and distribution of CO dehydrogenase structural genes in
 carboxydophilic bacteria.";
 RT Arch. Microbiol. 152:335-341(1989).
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + FERROCYTOCHROME B-561 - CO(2) +
 CC 2 H(+) + FERROCYTOCHROME B-561.
 CC -1- CORFACTOR: MOLYBDENUM.
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 CC SMALL.
 DR PIR: P0144; P0144.
 DR Oxidoreductase; Molybdenum.
 FT NON_TER 21
 FT NON_TER 21
 SQ SEQUENCE 21 AA; 2270 MW; 68D4380629401B9C CRC64;

Query Match 19.4%; Score 21; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2e+03;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 GHP 11
 DB 12 GHP 14

RESULT 19
 NSK1_SARBU STANDARD; PRT; 9 AA.
 ID NSK1_SARBU
 AC P41492;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE NEOSULFAKININ-I (NEB-SK-I).
 OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Oestrideae; Sarcophagidae; Sarcophaga.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-HEAD;
 RX MEDLINE-93083101; PubMed-1360367;
 RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.;
 RT Isolation and primary structure of two sulfakinin-like peptides from
 RT the fleshfly, Neobellieria bullata.";
 RL Comp. Biochem. Physiol. 103C:135-142(1992).
 CC -1- FUNCTION: MYOTROPIC PEPTIDE.
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECISTOKININ FAMILY.
 DR INTERPRO: IPR001651; -
 DR PROSITE: PS00259; GASTRIN; 1.
 KW Neuropeptide; Amidation; Sulfation.
 FT MOD_RES 4
 FT MOD_RES 4
 FT MOD_RES 4
 SQ SEQUENCE 9 AA; 1187 MW; 8B0A0691B865A0A CRC64;

Query Match 18.5%; Score 20; DB 1; Length 9;
 Best Local Similarity 60.0%; Pred. No. 8.9e+04;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 9 GHPRY 13
 DB 5 GHPRF 9

RESULT 20
 LSK2_LEUMA STANDARD; PRT; 10 AA.
 ID LSK2_LEUMA
 AC P09039;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE LEUCOSULFAKININ-II (LSK-II).
 OS Leucophaea maderae (Madeira cockroach), and
 OS Periplaneta americana (American cockroach).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberoidea; Blaberidae; Leucophaea.
 RN [1]
 RP SEQUENCE.
 RC SPECIES=L.MADERAE;
 RX MEDLINE=87048769; PubMed=3778455;
 RA Nachman R.J., Holman G.M., Cook B.J., Haddon W.F., Ling N.;
 RT "Leucosulfakinin-II, a blocked sulfated insect neuropeptide with
 RT homology to cholecystokinin and gastrin."
 RL Biochem. Biophys. Res. Commun. 140:357-364 (1986).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=P.AMERICANA; TISSUE=CORPORA CARDIACA;
 RX MEDLINE=90137190; PubMed=2615921;
 RA Veenstra J.A.;
 RT "Isolation and structure of two gastrin/CCK-like neuropeptides from
 RT the American cockroach homologous to the leucosulfakinins."
 RL Neuropeptides 14:145-149 (1989).
 CC -1- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF
 CC THE COCKROACH HINDGUT. STIMULATES MUSCLE CONTRACTION OF HINDGUT.
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 DR PIR: A26335; GMR02.
 DR INTERPRO: IPR001651;
 DR PROSITE: PS00259; GASTRIN: 1.
 KW Hormone; Amidation; Sulfatation.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 5 5 SULFATATION (IN L.MADERAE, BUT NOT IN
 FT AMIDATION).
 FT MOD_RES 10 10 AMIDATION.
 FT SEQUENCE 10 AA; 1255 MW; 9BAF5391E8B5AAA CRC64;
 QY 9 GHPRY 13
 DB 6 GHMRF 10
 Query Match 18.5%; Score 20; DB 1; Length 10;
 Best Local Similarity 60.0%; Pred. No. 1.3e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 RESULT 21
 LSK1_LEUMA STANDARD; PRT; 11 AA.
 AC P04428;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE LEUCOSULFAKININ-I (LSK-I).
 OS Leucophaea maderae (Madelira cockroach).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberoidea; Blaberidae; Leucophaea.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=86315858; PubMed=3749893;
 RA Nachman R.J., Holman G.M., Haddon W.F., Ling N.;
 RT "Leucosulfakinin, a sulfated insect neuropeptide with homology to
 RT gastrin and cholecystokinin."
 RL Science 234:71-73 (1988).
 CC -1- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF
 CC THE COCKROACH HINDGUT. INHIBITS MUSCLE CONTRACTION OF HINDGUT.
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 DR INTERPRO: IPR001651;
 DR PROSITE: PS00259; GASTRIN: 1.
 KW Hormone; Amidation; Sulfatation.
 FT MOD_RES 6 6 SULFATATION.
 FT MOD_RES 11 11 AMIDATION.
 FT SEQUENCE 11 AA; 1459 MW; 7EA06080E8B5AAB CRC64;

Query Match 18.5%; Score 20; DB 1; Length 11;
 Best Local Similarity 60.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 9 GHPRY 13
 DB 7 GHMRF 11
 Query Match 18.5%; Score 20; DB 1; Length 11;
 Best Local Similarity 60.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 RESULT 22
 LSKP_PERAM STANDARD; PRT; 11 AA.
 AC P36885;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE PERISULFAKININ (PER-SK-I).
 OS Periplaneta americana (American cockroach).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blattodea; Blattidae; Periplaneta.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=CORPORA CARDIACA;
 RX MEDLINE=90137190; PubMed=2615921;
 RA Veenstra J.A.;
 RT "Isolation and structure of two gastrin/CCK-like neuropeptides from
 RT the American cockroach homologous to the leucosulfakinins."
 RL Neuropeptides 14:145-149 (1989).
 CC -1- FUNCTION: STIMULATES HINDGUT CONTRACTIONS.
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 DR PIR: A60656; A60656.
 DR INTERPRO: IPR001651;
 DR PROSITE: PS00259; GASTRIN: 1.
 KW Hormone; Amidation; Sulfatation.
 FT MOD_RES 6 6 SULFATATION.
 FT MOD_RES 11 11 AMIDATION.
 FT SEQUENCE 11 AA; 1445 MW; 8BA06080E8B5AAA CRC64;
 QY 9 GHPRY 13
 DB 7 GHMRF 11
 Query Match 18.5%; Score 20; DB 1; Length 11;
 Best Local Similarity 60.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 RESULT 23
 LSKR_LOCOMI STANDARD; PRT; 12 AA.
 AC P47733;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE SULFAKININ (LOW-SK).
 OS Locusta migratoria (Migratory locust).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Caelifera;
 OC Acridoidea; Acrididae; Locusta.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=BRAIN;
 RX Schoofs L., Holman G.L., Hayes T.K., Nachman R.J., de Loof A.;
 RT (In) Mc Caffery A. "Wilson I. (eds.)";
 RT Chromatography and Isolation of Insect Hormones and Pheromones,
 RL pp. 231-241, Plenum Press, New York (1990).
 CC -1- FUNCTION: MYOTROPIC PEPTIDE.
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 DR INTERPRO: IPR001651;
 DR PROSITE: PS00259; GASTRIN: 1.

KW Hormone: Amidation; Sulfatation.
 FT MOD.RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD.RES 7 7 SULFATATION (POTENTIAL).
 FT MOD.RES 12 12 AMIDATION.
 SQ SEQUENCE 12 AA; 1440 MW; 9B5BDA9BD6B5A5A CRC64;

Query Match 18.5%; Score 20; DB 1; Length 12;
 Best Local Similarity 60.0%; Pred. No. 1.6e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 9 GHPRY 13
 1 1
 1 1
 DB 8 GHMRF 12

RESULT 24
 NSR2_SARBU STANDARD; PRT; 14 AA.
 ID NSR2_SARBU
 AC P41493;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE NEOSULFAKININ-II (NEB-SK-II).
 OS Sarcophaga bullata (grey flesh fly) (Neobellieria bullata).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Oestroidea; Sarcophagidae; Sarcophaga.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-HEAD;
 RX MEDLINE-93083101; PubMed-1360367;
 RA Fonagy A., Schoofs L., Proost P., Van Damme J., de Loof A.;
 RT "Isolation and primary structure of two sulfakinin-like peptides from
 the fleshfly, Neobellieria bullata."
 RL Comp. Biochem. Physiol. 103C:135-142(1992).
 CC -1- FUNCTION: MYOTROPIC PEPTIDE.
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 DR INTERPRO; IPRO01651; -;
 DR PROSITE; PS00259; GASTRIN; 1.
 KW Neuropeptide; Amidation; Sulfatation.
 FT MOD.RES 9 9 SULFATATION (POTENTIAL).
 FT MOD.RES 14 14 AMIDATION (POTENTIAL).
 SQ SEQUENCE 14 AA; 1794 MW; 8B4E06D5B61C62AA CRC64;

Query Match 18.5%; Score 20; DB 1; Length 14;
 Best Local Similarity 60.0%; Pred. No. 1.9e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 9 GHPRY 13
 1 1
 1 1
 DB 10 GHMRF 14

RESULT 25
 UC08_MAIZE STANDARD; PRT; 15 AA.
 ID UC08_MAIZE
 AC P80614;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 159)
 DE (FRAGMENT).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-COLEOPTILE;
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
 RA Perrotlet J.-C., Zivy M., de Vienne D.;
 RT "The maize two dimensional gel protein database: towards an integrated

RT genome analysis program";
 RL Theor. Appl. Genet. 93:997-1005(1996).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.4, ITS MW IS: 38.8 KDA.
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY.
 DR MAIZE-2DPAGE; P80614; COLEOPTILE.
 DR MAIZE2DB; 123934; -;
 FT NON_TER 1 1
 FT NON_TER 15 15
 SQ SEQUENCE 15 AA; 1785 MW; 1978B1D6A84D8D CRC64;

Query Match 18.5%; Score 20; DB 1; Length 15;
 Best Local Similarity 42.9%; Pred. No. 2.1e+03;
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 YAIKTH 10
 1 1
 1 1
 DB 8 YFVYPGH 14

Search completed: February 5, 2001, 10:56:01
 Job time: 501 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:55:24 ; Search time 93.77 Seconds
(without alignments)
24.999 Million cell updates/sec

Title: US-08-981-824-4

Sequence: 1 TLKYAIKTHPRYFNQLSTG 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MNC:*
8: SP_ORGANELLE:*
9: SP_PHASE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	29.6	21	4	Q9UCAT7
2	30.5	28.2	23	10	Q9S821
3	30	27.8	25	2	Q26056
4	29.5	27.3	20	5	Q9TWM2
5	29	26.9	24	2	Q9RS1
6	29	26.9	24	4	Q15133
7	28	25.9	19	2	Q53502
8	28	25.9	20	2	Q9RS00
9	28	25.9	25	2	Q9RS50
10	27	25.0	12	4	Q16452
11	26.5	24.1	19	6	Q94149
12	26	24.1	19	6	Q9TR84
13	26	24.1	20	3	Q13594
14	26	24.1	21	6	Q9TR06
15	25	23.1	10	12	Q85462
16	25	23.1	15	7	Q9TNO1
17	25	23.1	18	10	P82242
18	25	23.1	21	12	Q93191
19	25	23.1	23	2	Q92E34

Result No.	Score	Query Match	Length	DB ID	Description
20	25	23.1	25	11	Q9TK04
21	24	22.2	8	11	Q9JLD7
22	24	22.2	15	2	Q52135
23	24	22.2	22	1	Q9UWK5
24	24	22.2	22	4	Q9UCK8
25	23	21.3	13	11	Q9QW04
26	23	21.3	15	8	Q9S586
27	23	21.3	15	12	Q9PXC5
28	23	21.3	16	10	Q9S8A0
29	23	21.3	16	12	Q90049
30	23	21.3	18	10	Q9S7E5
31	23	21.3	18	12	Q41588
32	23	21.3	21	11	Q62538
33	23	21.3	22	13	Q9PRN2
34	23	21.3	23	4	Q9UCB0
35	23	21.3	23	13	Q13029
36	23	21.3	25	5	Q26087
37	23	21.3	25	12	Q84040
38	22	20.4	13	10	Q43174
39	22	20.4	15	8	Q9S751
40	22	20.4	16	4	Q14495
41	22	20.4	16	6	Q9TR11
42	22	20.4	17	2	Q52210
43	22	20.4	19	8	Q9TMN2
44	22	20.4	19	8	Q9TMN1
45	22	20.4	19	8	Q9T304

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	32	29.6	21	4	Q9UCAT7
2	30.5	28.2	23	10	Q9S821
3	30	27.8	25	2	Q26056
4	29.5	27.3	20	5	Q9TWM2
5	29	26.9	24	2	Q9RS1
6	29	26.9	24	4	Q15133
7	28	25.9	19	2	Q53502
8	28	25.9	20	2	Q9RS00
9	28	25.9	25	2	Q9RS50
10	27	25.0	12	4	Q16452
11	26.5	24.1	19	6	Q94149
12	26	24.1	19	6	Q9TR84
13	26	24.1	20	3	Q13594
14	26	24.1	21	6	Q9TR06
15	25	23.1	10	12	Q85462
16	25	23.1	15	7	Q9TNO1
17	25	23.1	18	10	P82242
18	25	23.1	21	12	Q93191
19	25	23.1	23	2	Q92E34

Query Match	Score	DB 4	Length	Best Local Similarity	Pred. No.	1.5e+02	Matches	5	Conservative	0	Mismatches	2	Indels	0	Gaps	0
Q9UCAT7	32	29.6	21	71.4%	32	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q9S821	30.5	28.2	23	71.4%	30.5	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q26056	30	27.8	25	71.4%	30	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q9TWM2	29.5	27.3	20	71.4%	29.5	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q9RS1	29	26.9	24	71.4%	29	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q15133	29	26.9	24	71.4%	29	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q53502	28	25.9	19	71.4%	28	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q9RS00	28	25.9	20	71.4%	28	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q9RS50	28	25.9	25	71.4%	28	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q16452	27	25.0	12	71.4%	27	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q94149	26.5	24.1	19	71.4%	26.5	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q9TR84	26	24.1	19	71.4%	26	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q13594	26	24.1	20	71.4%	26	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q9TR06	26	24.1	21	71.4%	26	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q85462	25	23.1	10	71.4%	25	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q9TNO1	25	23.1	15	71.4%	25	1.5e+02	5	5	5	0	0	0	0	0	0	0
P82242	25	23.1	18	71.4%	25	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q93191	25	23.1	21	71.4%	25	1.5e+02	5	5	5	0	0	0	0	0	0	0
Q92E34	25	23.1	23	71.4%	25	1.5e+02	5	5	5	0	0	0	0	0	0	0

RESULT	6		
015133			
ID	015133;	PRELIMINARY;	PRT; 24 AA.
AC	015133;		
DT	01-NOV-1996 (TREMBLrel. 01.	Created)	
DT	01-NOV-1996 (TREMBLrel. 01.	Last sequence update)	
DT	01-NOV-1998 (TREMBLrel. 08.	Last annotation update)	
DE	GHF-1 TRANSCRIPTION FACTOR (FRAGMENT).		
GN	GHF-1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=PLACENTA;		
RX	MEDLINE=95237623; PubMed=7721104;		
RA	Delhase M., Valla V., Hooghe-Peters E.L., Castriello J.L.,		
RT	A novel placental transcription factor is produced by alternativ		

RT splicing of the human GHF-1/PIT-1 gene."
 RL Gene 155:273-275(1995).
 DR EMBL: X77223; CAAS4441.1;
 DR EMBL: X77224; CAAS4441.1; JOINED.
 FT NON_TER 1
 SQ SEQUENCE 24 AA; 2571 MW; EC12619EC5E28C07 CRC64;

Query Match 26.9%; Score 29; DB 4; Length 24;
 Best Local Similarity 44.4%; Pred. No. 5.2e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 2 LKVAIKTGHPRYFNQST 19
 Db 4 LKVSVPSCN--YGNQPSI 19

RESULT 7
 ID Q53502 PRELIMINARY; PRT; 19 AA.
 AC Q53502;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE ORF PROTEIN (FRAGMENT).
 OS Lactobacillus paracasei.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 OC Lactobacillus.
 NX NCBI_TaxID=1597;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95219756; Pubmed=7704831;
 RA Djordjevic G., Bojovic B., Bantina A., Topisirovic L.;
 RT "Cloning of promoter-like sequences from Lactobacillus paracasei
 subsp. paracasei CG1 and their expression in Escherichia coli."
 RT Lactococcus lactis and Lactobacillus reuteri."
 RL Can. J. Microbiol. 40:1043-1050(1994).
 DR EMBL: S76788; AAB3947.2; -.
 FT NON_TER 19
 SQ SEQUENCE 19 AA; 2135 MW; 41BCAAC6CC513BA5 CRC64;

Query Match 25.9%; Score 28; DB 2; Length 19;
 Best Local Similarity 71.4%; Pred. No. 5.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKVAIKT 8
 Db 1 MKVAIKS 7

RESULT 8
 ID Q9R500 PRELIMINARY; PRT; 20 AA.
 AC Q9R500;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE METALLOENDOPEPTIDASE I, METALLOENDOPEPTIDASE II, SGMPI,
 DE SGMPI-ZINC-ENDOPEPTIDASE.
 OS Streptomyces griseus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 NX NCBI_TaxID=1911;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=9212147; Pubmed=1769959;
 RA Tsuyuki H., Kajiwara K., Fujita A., Kumazaki T., Ishii S.;
 RT "Purification and characterization of Streptomyces griseus
 metalloendopeptidases I and II."
 RL J. Biochem. 110:339-344(1991).
 SQ SEQUENCE 20 AA; 2030 MW; 27929F80B6299162 CRC64;

Query Match 25.9%; Score 28; DB 2; Length 20;
 Best Local Similarity 38.5%; Pred. No. 6.2e+02;
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 8 TGHPRYFNQSTG 20
 Db 2 TGNTRYNGQVTLG 14

RESULT 9
 ID Q9R500 PRELIMINARY; PRT; 25 AA.
 AC Q9R500;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE PERICILLIN-BINDING PROTEIN 4, PPB4.
 OS Enterococcus hirae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
 OC Enterococcus.
 NX NCBI_TaxID=1354;
 RX [1]
 RP SEQUENCE.
 RA MEDLINE=92038914; Pubmed=1936941;
 RA Jacques P., el Kharrubi A., Van Beeumen J., Piras G., Coyette J.,
 RA Ghysen J.M.;
 RT "Mode of membrane insertion and sequence of a 32-amino acid peptide
 stretch of the penicillin-binding protein 4 of Enterococcus hirae."
 RL FEMS Microbiol. Lett. 66:119-123(1991).
 SQ SEQUENCE 25 AA; 2707 MW; 9189FB2F034C44BF CRC64;

Query Match 25.9%; Score 28; DB 2; Length 25;
 Best Local Similarity 50.0%; Pred. No. 7.9e+02;
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLKVAIKTGHPR 12
 Db 10 TLTLTIQTSPXR 21

RESULT 10
 ID Q16452 PRELIMINARY; PRT; 12 AA.
 AC Q16452;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 08, Last annotation update)
 DE COLLAGEN ALPHA 2(I) CHAIN CB PEPTIDE ALPHA 2CB4.
 GN COL1A2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92201825; Pubmed=1551666;
 RA Nicholas A.C., Oliver J., Renouf D.V., Heath D.A., Pope F.M.;
 RT "The molecular defect in a family with mild atypical osteogenesis
 imperfecta and extreme joint hypermobility: exon skipping caused by an
 RT 11-bp deletion from an intron in one COL1A2 allele."
 RL Hum. Genet. 88:627-633(1992).
 DR EMBL: S89896; AAB21862.1; -.
 SQ SEQUENCE 12 AA; 1161 MW; D5BD7D21FAE861B6 CRC64;

Query Match 25.0%; Score 27; DB 4; Length 12;
 Best Local Similarity 80.0%; Pred. No. 5.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 KTGHPR 11
 Db 11 KTGHPR 11

Db 5 Q7GHP 9

RESULT 11

094149 PRELIMINARY; PRT; 25 AA.

AC 094149; PRELIMINARY; PRT; 25 AA.
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE PROBABLY NOT EXPRESSED DUE TO SNRNA OVERLAPPING (FRAGMENT).
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomycetes.
 RN NCBI_TaxID=4932;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Boyer J., Fairhead C., Gallon L., Gallsion F., Michaux G.,
 RA Thierry A., Dujon B.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA MIPS:
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 275144; CAA94457.1;
 FT NON_TER 1 1
 SQ SEQUENCE 25 AA; 3023 MW; D6C9225838A38D04 CRC64;

Query Match 24.5%; Score 26.5; DB 3; Length 25;
 Best Local Similarity 31.6%; Pred. No. 1.4e+03;
 Matches 6; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 2 LKYAIKTGHPRYFNOLSTG 20
 DB 4 LQYVI---YPKHFTKKPG 19

RESULT 12

09TRR4 PRELIMINARY; PRT; 19 AA.

AC 09TRR4; PRELIMINARY; PRT; 19 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 13, Last annotation update)
 DE CALYCULIN-ASSOCIATED PROTEIN, CAP50-CA2+/PHOSPHOLIPID-BINDING PROTEIN
 DE L-21 FRAGMENT.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 RN NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=92250478; PubMed=1533622;
 RA Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
 RA "A calyculin-associated protein is a newly identified member of the
 RT CA2+/phospholipid-binding proteins, annexin family."
 RL J. Biol. Chem. 267:8919-8924(1992).
 DR HSSP; P79134; IAVC.
 SQ SEQUENCE 19 AA; 2186 MW; 576DC1604E19BCB8 CRC64;

Query Match 24.1%; Score 26; DB 6; Length 19;
 Best Local Similarity 35.3%; Pred. No. 1.2e+03;
 Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLKYAIKTGHPRYFNOL 17
 DB 1 TLEBAIRSDTSGHFORU 17

RESULT 13

013594

ID 013594 PRELIMINARY; PRT; 20 AA.

AC 013594; PRELIMINARY; PRT; 20 AA.
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE ORF YOR222M (FRAGMENT).
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomycetes.
 RN NCBI_TaxID=4932;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Boyer J., Fairhead C., Gallon L., Gallsion F., Michaux G.,
 RA Thierry A., Dujon B.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS:
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 275131; CAA94442.2;
 FT NON_TER 1 1
 SQ SEQUENCE 20 AA; 2420 MW; 6685F6B03B1E695 CRC64;

Query Match 24.1%; Score 26; DB 3; Length 20;
 Best Local Similarity 31.6%; Pred. No. 1.3e+03;
 Matches 6; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 2 LKYAIKTGHPRYFNOLSTG 20
 DB 1 LMLVFTGMNFFRDLKYG 19

RESULT 14

09TRO6 PRELIMINARY; PRT; 21 AA.

AC 09TRO6; PRELIMINARY; PRT; 21 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CALYCULIN-ASSOCIATED PROTEIN PEPTIDE L-13, CAP-50-ANNEXIN.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 RN NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=92317074; PubMed=1618851;
 RA Mizutani A., Usuda N., Tokumitsu H., Minami H., Yasui K.,
 RA Kobayashi R., Hidaka H.;
 RA "CAP-50, a newly identified annexin, localizes in nuclei of cultured
 RT fibroblast 3Y1 cells."
 RL J. Biol. Chem. 267:13498-13504(1992).
 SQ SEQUENCE 21 AA; 2352 MW; 998D0E8AC173DA19 CRC64;

Query Match 24.1%; Score 26; DB 6; Length 21;
 Best Local Similarity 41.2%; Pred. No. 1.4e+03;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLKYAIKTGHPRYFNOL 17
 DB 1 TLEBAIRSDTSGHFOOL 17

RESULT 15

085462 PRELIMINARY; PRT; 10 AA.

AC 085462; PRELIMINARY; PRT; 10 AA.
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
 DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)


```

DE ENV PROTEIN (FRAGMENT).
GN ENV.
OS Avian sarcoma virus.
OS Viruses; Retrov. Viruses; Retroviridae; Avian type C retroviruses.
OX NCBI_TaxID=11876;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=88275035; PubMed=2839694;
RX Katz R.A., Kotler M., Skalka A.M.;
RT "cis-acting intron mutations that affect the efficiency of avian
retroviral RNA splicing: implication for mechanisms of control.";
RL J. Virol. 62:2686-2695(1988).
DR EMBL: M21395; AAA42413.1; -.
FT NON_TER 1 10
FT SEQUENCE 10 AA; 1007 MW; 5F751D67B5861A7 CRC64;

Query Match 23.1%; Score 25; DB 12; Length 10;
Best Local Similarity 57.1%; Pred. NO. 8.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 AKTGHP 11
DB 1 AVLTGTP 7

RESULT 16
O9TNO1 PRELIMINARY; PRT; 15 AA.
AC O9TNO1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE BETA 2M-CLASS I-BINDING PEPTIDE-MAJOR HISTOCOMPATIBILITY COMPLEX
DE H-2K-B-SPECIFIC MOLECULE POORLY ASSOCIATED WITH BETA 2-MICROGLOBULIN.
Kus SP.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE.
RX MEDLINE=94240094; PubMed=8183884;
RA Joyce S., Kuzushima K., Repecs G., Angeletti R.H., Nathanson S.G.;
RT "Characterization of an incompletely assembled major
histocompatibility class I molecule (H-2Kb) associated with unusually
RT long peptides: implications for antigen processing and presentation.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4145-4149(1994).
KM MHC.
SQ SEQUENCE 15 AA; 1754 MW; 2CEACDA0F1180DC7 CRC64;

Query Match 23.1%; Score 25; DB 7; Length 15;
Best Local Similarity 71.4%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 11 PRYFNOI 17
DB 1 PRYLLOI 7

RESULT 17
P82242 PRELIMINARY; PRT; 18 AA.
AC P82242;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE MAJOR POLLEN ALLERGEN PLA 1 (FRAGMENT).
OC Plantago lanceolata (narrow-leaved plantain).
OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OX Lamiales; Plantaginaceae; Plantago.

```

```

OX NCBI_TaxID=39414;
RN [1]
RP SEQUENCE.
RC TISSUE-POLLEN;
RA Calabozo B., Barber D., Polo F.;
RT "Purification and characterization of the main allergen of Plantago
lanceolata pollen, Pla 1 I.";
RL J. Allergy Clin. Immunol. 0:0-0(1999).
KM Allergen.
FT NON_TER 18 18
FT SEQUENCE 18 AA; 2054 MW; BD780B5758B53559D CRC64;

Query Match 23.1%; Score 25; DB 10; Length 18;
Best Local Similarity 44.4%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 7 KITGHPRYFN 15
DB 2 QTSHPAFH 10

RESULT 18
O93191 PRELIMINARY; PRT; 21 AA.
ID O93191;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE ORF8.
OS Porcine circovirus.
OC Viruses; ssDNA viruses; Circoviridae; Circovirus.
OX NCBI_TaxID=46221;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PORCINE CIRCOVIRUS TYPE II;
RX MEDLINE=98418498; PubMed=9747726;
RA Meenan B.M., McNeill F.M., Todd D., Kennedy S., Jewhurst V.,
RA Ellis J.A., Haasard L.E., Clark E.G., Haines D.M., Allen G.M.;
RT "Characterization of novel circovirus DNAs associated with wasting
RT syndromes in pigs.";
RL J. Gen. Virol. 79:2171-2199(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-PORCINE CIRCOVIRUS TYPE II;
RA Meenan B.M.;
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF055393; AAC35324.1; -.
DR EMBL: AF055393; AAC35324.1; -.
SQ SEQUENCE 21 AA; 2369 MW; BAB8F6E2184C1013 CRC64;

Query Match 23.1%; Score 25; DB 12; Length 21;
Best Local Similarity 27.3%; Pred. No. 2e+03;
Matches 3; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 LKVAIKTGHP 12
DB 3 IDHTVSVDHPR 13

RESULT 19
O9ZECJ4 PRELIMINARY; PRT; 23 AA.
ID O9ZECJ4;
AC O9ZECJ4;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE DNAJ2 PROTEIN (FRAGMENT).
GN DNAJ2.
OS Anabaena sp.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=1167;

```

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PCC 7120;
 RA Pohl B.;
 RL Thesis (1999), University of Bonn, Botanical Institute.
 DR EMBL: AJ132709; CA10746.1; -.
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2493 MW; AC5084286BC591ED CRC64;

Query Match 23.1%; Score 25; DB 2; Length 23;
 Best Local Similarity 36.4%; Pred. No. 2.2e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 7 KTGHPRYFNOL 17
 DB 3 KSEPTYSLL 13

RESULT 20
 ID 09JRK04 PRELIMINARY; PRT; 25 AA.
 AC 09JRK04;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE FRUCTOSE-1,6-BISPHOSPHATASE (EC 3.1.3.11) (FRAGMENT).
 GN FBPA5E 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tilmann H., Stein S., Liehr T., Eschrich K.;
 RT "Structure and chromosomal localization of the human and mouse muscle
 fructose-1,6-bisphosphatase genes.";
 RL Gene 247:241-253(2000).
 DR EMBL: AJ243025; CAB90672.1; -.
 KW Hydrolyase.
 FT NON_TER 25
 FT NON_TER 25
 SQ SEQUENCE 25 AA; 2898 MW; E4A7A6F50F3A31D3 CRC64;

Query Match 23.1%; Score 25; DB 11; Length 25;
 Best Local Similarity 27.3%; Pred. No. 2.4e+03;
 Matches 3; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 5 AIKTGHPRYFN 15
 DB 1 SLNEGAKYFD 11

RESULT 21
 ID 09JLD7 PRELIMINARY; PRT; 8 AA.
 AC 09JLD7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE P53 TUMOR SUPPRESSOR (FRAGMENT).
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-KIDNEY.
 RC Laverdiere M., Beaudoin J., Lavigne A.;
 RT "Species-specific regulation of alternative splicing in the C-terminal
 region of the p53 tumor suppressor gene.";

RL Nucleic Acids Res. 28:1489-1497(2000).
 DR EMBL: AF190271; AAF43279.1; -.
 FT NON_TER 1
 SQ SEQUENCE 8 AA; 969 MW; 1DB6D9CDA1761E7 CRC64;

Query Match 22.2%; Score 24; DB 11; Length 8;
 Best Local Similarity 80.0%; Pred. No. 3.7e+05;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 10 HPRYF 14
 DB 2 HPRAF 6

RESULT 22
 ID 052135 PRELIMINARY; PRT; 15 AA.
 AC 052135;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE PLASMIN DAD1 SEX PHEROMONE INHIBITOR (IAD1) DETERMINANT.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Plasmid PADI.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
 OC Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91261999; PubMed-2128961;
 RA Clewell D.B., Pontius L.T., An F.Y., Ike Y., Suzuki A., Nakayama J.;
 RT "Nucleotide sequence of the sex pheromone inhibitor (IAD1) determinant
 of Enterococcus faecalis conjugative plasmid PADI.";
 RL Plasmid 24:156-161(1990).
 DR EMBL: M62888; AAA98040.1; -.
 KW Plasmid.
 SQ SEQUENCE 15 AA; 1874 MW; 0D9D07E3079E3559 CRC64;

Query Match 22.2%; Score 24; DB 2; Length 15;
 Best Local Similarity 36.4%; Pred. No. 2e+03;
 Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 YAIKTGHPRYF 14
 DB 2 YTVHYIRPF 12

RESULT 23
 ID 09UWK5 PRELIMINARY; PRT; 22 AA.
 AC 09UWK5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE 6S KDA TUNGSTEN CONTAINING FORMYL METHANOFURAN DEHYDROGENASE
 (FRAGMENT).
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacterium.
 OX NCBI_TaxID=2166;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-94213570; PubMed-8161283;
 RA Bertman P.A., Schmitz R.A., Linder D., Thauer R.K.;
 RT "Tungstate can substitute for molybdate in sustaining growth of
 Methanobacterium thermoautotrophicum. Identification and
 characterization of a tungsten isoenzyme of formylmethanofuran
 dehydrogenase.";
 RL Arch. Microbiol. 161:220-228(1994).
 SQ SEQUENCE 22 AA; 2575 MW; EC412A8DC463A33 CRC64;

Search completed: February 5, 2001, 10:55:26
 Job time: 915 sec

Query Match 22.2%; Score 24; DB 1; Length 22;
 Best Local Similarity 50.0%; Pred. No. 3e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKVAIKTG 9
 DB 1 MEYIKNG 8

RESULT 24

Q9UCK8 PRELIMINARY; PRT; 22 AA.
 AC Q9UCK8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ASPARTYLGLUCOSAMINIDASE ALPHA SUBUNIT (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93111925; PubMed=1281977;
 RA Rip J.W., Coulter-Wackie M.B., Rupa C.A., Gordon B.A.;
 RT "Purification and structure of human liver aspartylglucosaminidase.";
 RL Biochem. J. 288:1005-1010(1992).
 DR HSP; P20933; IAPY.
 SQ SEQUENCE 22 AA; 2179 MW; 9628EC8CFBDA28F CRC64;

Query Match 22.2%; Score 24; DB 4; Length 22;
 Best Local Similarity 100.0%; Pred. No. 3e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 KTGH 10
 DB 9 KTGH 12

RESULT 25

Q9QW04 PRELIMINARY; PRT; 13 AA.
 AC Q9QW04;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Pl.
 OS Mus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94043690; PubMed=7693750;
 RA Barnett L.A., Whitton J.L., Wada Y., Fujinami R.S.;
 RT "Enhancement of autoimmune disease using recombinant vaccinia virus
 encoding myelin proteolipid protein.";
 RL J. Neuroimmunol. 48:120-120(1993).
 SQ SEQUENCE 13 AA; 1538 MW; 2A33FF99B5807733 CRC64;

Query Match 21.3%; Score 23; DB 11; Length 13;
 Best Local Similarity 66.7%; Pred. No. 2.4e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 GHPRYF 14
 DB 8 GHPDKF 13

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:47:31 ; Search time 68.99 Seconds
(without alignments)
9.913 Million cell updates/sec

Title: US-08-981-824-4

Perfect score: 108
Sequence: 1 TLKXAIKTHPRFYNQJSTG 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.36.*
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	20	18 W18845	65 KD Glutamic ac1
2	108	100.0	20	18 W01796	Human 65 KD glutam
3	108	100.0	20	18 W12404	GAD65 residues 173
4	72	66.7	20	16 R72272	Glutamic acid deca
5	72	66.7	20	21 Y59572	GAD65 fragment, pe
6	60	55.6	20	16 R72271	Glutamic acid deca
7	60	55.6	20	21 Y59571	GAD65 fragment, pe
8	55	50.9	20	18 W18846	65 KD Glutamic ac1
9	55	50.9	20	18 W01797	Human 65 KD glutam
10	48	44.4	14	16 R76657	Peptide derived fr
11	48	44.4	14	18 W18865	65 KD Glutamic ac1
12	48	44.4	15	18 W12405	GAD65 residues 177

13	34	31.5	24	20 Y36566	Fragment of human
14	32	29.6	14	15 R49436	Calgranulin B posi
15	32	29.6	24	15 R49435	Calgranulin B posi
16	31	28.7	20	21 Y91936	HSRec2 peptide res
17	30	27.8	11	21 Y57311	P. aeruginosa cera
18	30	27.8	16	14 R43130	Rhinoviral HRV2 2A
19	30	27.8	16	18 W35218	Dilatoromer pepit
20	30	27.8	16	18 W35210	Dilatoromer pepit
21	30	27.8	16	19 W82934	Antipathogenic pep
22	30	27.8	16	19 W82926	Antipathogenic pep
23	30	27.8	17	19 W80790	Peptide 6 from D.
24	29	26.9	9	20 Y01008	Bacterial immune
25	29	26.9	15	20 W92558	Mouse beta-actin a
26	29	26.9	15	20 W92559	Mouse beta-actin a
27	29	26.9	22	21 Y58198	Human STRAP-1 pept
28	28	25.9	13	19 W70635	Second generation
29	28	25.9	13	21 Y99311	HLA class II bindi
30	28	25.9	15	16 R68754	Cytotoxic T lympho
31	28	25.9	15	16 R68755	Diabetes type I au
32	28	25.9	15	20 Y13251	Naturally occurin
33	28	25.9	15	20 Y13300	Naturally occurin
34	28	25.9	15	20 Y13277	Naturally occurin
35	28	25.9	16	15 R57905	Randomly generated
36	28	25.9	16	17 R90371	Residues 10-25 of
37	28	25.9	20	20 W74168	HJ loop peptide CA
38	28	25.9	23	13 R20236	NAF(3-25) peptide
39	28	25.9	25	16 R74073	Superantigen pept
40	28	25.9	25	20 Y39991	HIV Negative facto
41	28	25.9	25	20 Y26613	HIV-derived lipope
42	27.5	25.5	24	7 P60962	Polypeptide which
43	27	25.0	9	15 R62231	Toxoplasma gondii
44	27	25.0	9	16 R79703	Egf receptor Tyr k
45	27	25.0	9	21 Y67904	Egf receptor autop

ALIGNMENTS

RESULT 1
W18845
ID W18845 standard; peptide; 20 AA.
XX
AC W18845;
XX
DT 05-JAN-1998 (first entry)
XX
DE 65 KD Glutamic acid decarboxylase peptide fragment IV.
XX
KW GAD, 65 KD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
KW Insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
KW predisposition; autoimmune; tumour; rheumatoid arthritis;
KW multiple sclerosis.
XX
OS Synthetic.
XX
PN DE19526561-A1.
XX
PD 23-JAN-1997.
XX
PF 20-JUL-1995; 95DE-1026561.
XX
PR 20-JUL-1995; 95DE-1026561.
XX
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
XX
PI Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;
PI Pozzilli P, Stahl P;
DR WPI; 1997-088254/09.
XX
PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
XX Involving intradermal admin. of auto-reactive substances

PS Claim 11; Page 9; 12pp; German.
 CC W18842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predisposition to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 CC
 CC Sequence 20 AA;
 CC
 CC Query Match 100.0%; Score 108; DB 18; Length 20;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e-12;
 CC Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Oy 1 TLKYAIKTGHPRYFNQLSTG 20
 CC |||||
 CC Db 1 tlkyaiktghpryfnqlstg 20
 CC
 CC RESULT 2
 CC W01796
 CC ID W01796 standard; peptide; 20 AA.
 CC XX
 CC AC W01796;
 CC XX
 CC DT 15-OCT-1997 (first entry)
 CC XX
 CC DE Human 65 kD glutamine decarboxylase peptide.
 CC XX
 CC KW Human; glutamine decarboxylase; GAD; diagnosis; predisposition;
 CC KW tumour; immunological; disease; autoimmune; diabetes; reagent;
 CC KW determination; T cell; subpopulation; medication; treatment;
 CC KW prevention; production; antigen; immunogen; tolerogen; isolation;
 CC KW reinjection; inactivation.
 CC XX
 CC OS Homo sapiens.
 CC XX
 CC PN DE19525784-A1.
 CC XX
 CC PD 16-JAN-1997.
 CC XX
 CC PF 14-JUL-1995; 95DE-1025784.
 CC XX
 CC PR 14-JUL-1995; 95DE-1025784.
 CC XX
 CC PA (BOE) BOEHRINGER MANNHEIM GMBH;
 CC XX
 CC PI Albert W. Bollard C. Endl J. Jung G. Schendel D;
 CC XX
 CC PT Stahl P, Van Endert P;
 CC XX
 CC PS WPI: 1997-078452/08.
 CC
 CC Claim 1; Page 12; 15pp; German.
 CC
 CC The present peptide is a fragment of the human 65 kD glutamine
 CC decarboxylase (GAD), which can be used to diagnose, or diagnose a
 CC predisposition to, a tumour or immunological disease, preferably an
 CC autoimmune disease, especially diabetes. It can also be used as a
 CC reagent to determine specific T cell subpopulations, in medicaments
 CC to treat or prevent immunological diseases, preferably autoimmune
 CC diseases, especially diabetes, to produce antigens, especially
 CC immunogens or tolerogens and to isolate specific T cell
 CC subpopulations, which can be used to produce antigens or for
 CC reinjection, optionally after inactivation.

XX Sequence 20 AA;
 CC
 CC Query Match 100.0%; Score 108; DB 18; Length 20;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e-12;
 CC Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Oy 1 TLKYAIKTGHPRYFNQLSTG 20
 CC |||||
 CC Db 1 tlkyaiktghpryfnqlstg 20
 CC
 CC RESULT 3
 CC W12404
 CC ID W12404 standard; peptide; 15 AA.
 CC XX
 CC AC W12404;
 CC XX
 CC DT 08-OCT-1997 (first entry)
 CC XX
 CC DE GAD65 residues 173-187.
 CC XX
 CC KW 65 kD glutamic acid decarboxylase; human; GAD65; enzyme; pancreas;
 CC KW neuron; central nervous system; type I diabetes; autoimmune response;
 CC KW T cell; therapy.
 CC XX
 CC OS Homo sapiens.
 CC XX
 CC FH Key Location/Qualifiers
 CC FT Misc-difference 1.15
 CC FT /note- "optionally substituted, providing at least 7
 CC FT residues remain wild type"
 CC XX
 CC PN W09700891-A1.
 CC XX
 CC PD 09-JAN-1997.
 CC XX
 CC PF 24-JUN-1996; 96WO-US10790.
 CC XX
 CC PR 23-JUN-1995; 95US-0494624.
 CC XX
 CC PA (KENN-) KENNEDY INST RHEUMATOLOGY.
 CC PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
 CC PA (SBR-) ST BARTHOLOMEW'S HOSPITAL CENT CLINICAL.
 CC XX
 CC PI Conlon PJ, Gaur A, Leslie RDG, Ling N, Londei M;
 CC XX
 CC DR WPI: 1997-087322/08.
 CC
 CC Claim 22; Page 7; 28pp; English.
 CC
 CC W12403-W12413 represent fragments and analogues of the 65 kD isoform
 CC of human glutamic acid decarboxylase (GAD65) (see W12402 for full length
 CC wild type protein). GAD is an enzyme expressed in the beta cells of the
 CC pancreas, and in neurons of the central nervous system. There are two
 CC isoforms of GAD, a 67 kD isoform, and GAD65. Immunodominant regions of
 CC GAD65 have been identified in Type I diabetic patients. These GAD65
 CC fragments, and analogues, are used in the methods of the invention. The
 CC methods are for detecting or treating diabetes or a predisposition to
 CC diabetes. The peptides can also be used for ameliorating an autoimmune
 CC response in a patient. Alteration of the native peptides with selective
 CC changes of crucial residues can induce unresponsiveness or change the
 CC responsiveness of antigen-specific autoreactive T cells. The peptide
 CC analogues compete for binding to MHC and do not cause proliferation of
 CC the corresponding native peptide-specific T cells.
 CC
 CC Sequence 15 AA;

Query Match 68.5%; Score 74; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TGHPRYFNQSTG 20
| | | | | | | | | | | | | | | | |
Db 1 tghpryfnqstg 13

RESULT 4

R72272
ID R72272 standard; Peptide; 20 AA.

AC R72272;

DT 13-NOV-1995 (first entry)

DE Glutamic acid decarboxylase (GAD65) fragment.

KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
KM Insulin-dependent diabetes mellitus; stiff man disease.

OS Homo sapiens.

PN W09507992-A.

PD 23-MAR-1995.

PF 24-AUG-1994; 94WO-US09478.

PR 17-SEP-1993; 93US-0123859.

XX (REGC) UNIV CALIFORNIA.

PI Clare-Saltzer MJ, Erlander MG, Kaufman DL, Tobin AJ;

DR WPI; 1995-131360/17.

PT New polypeptide fragments of glutamic acid decarboxylase - for

PT diagnosis and treatment of autoimmune disease, esp. insulin

PT antibodies, hybridoma(s) etc.

PS Example 11; Page 76; 100pp; English.

CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic

CC acid decarboxylase (GAD65) respectively, from which the GAD65

CC fragments described in R72261-R72298 were derived. These fragments

CC can be used to detect autoantibodies against GAD, e.g. to diagnose

CC and treat GAD-related autoimmune disorders, such as insulin

CC dependant diabetes mellitus or stiff man disease.

SO Sequence 20 AA;

QY 7 KTGHPRYFNQSTG 20
| | | | | | | | | | | | | | | | |
Db 1 kyghpryfnqstg 14

RESULT 5

ID Y59572 standard; peptide; 20 AA.

AC Y59572;

DT 03-APR-2000 (first entry)

DE GAD65 fragment, peptide #12.

KW GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;

KM Insulin dependent diabetes mellitus; Stiff man disease; diagnosis;

OS Homo sapiens.

PN US598366-A.

PD 07-DEC-1999.

PF 09-APR-1997; 97US-0827618.

PR 07-JUN-1995; 95US-0485725.

PR 21-SEP-1990; 90US-0586536.

PR 18-JUN-1991; 91US-0716909.

XX (REGC) UNIV CALIFORNIA.

PI Tobin AJ, Kaufman DL, Erlander MG;

DR WPI; 2000-095930/08.

PT Ameliorating glutamic acid decarboxylase associated autoimmune

PT disorders such as insulin dependent diabetes mellitus and Stiff man

PT disease -

PS Example 11; Column 42; 61pp; English.

CC This sequence represents a fragment of the glutamic acid decarboxylase 65

CC (GAD65) protein. The invention relates to a method of ameliorating GAD

CC associated autoimmune disorders by administering a GAD65 peptide to the

CC patient. The method can be used for ameliorating GAD associated

CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)

CC and Stiff man disease. GAD65 can also be useful for screening drugs that

CC alter GAD function, for generating monoclonal antibodies and in

CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM

CC and the diagnosis is quite easy. It is also possible to obtain much

CC larger quantities of polypeptide via recombinant techniques than are

CC available from natural sources.

SO Sequence 20 AA;

QY 7 KTGHPRYFNQSTG 20
| | | | | | | | | | | | | | | | |
Db 1 kyghpryfnqstg 14

DT 13-NOV-1995 (first entry)

DE Glutamic acid decarboxylase (GAD65) fragment.

KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;

KM Insulin-dependent diabetes mellitus; stiff man disease.

OS Homo sapiens.

PN W09507992-A.

PD 23-MAR-1995.

PF 24-AUG-1994; 94WO-US09478.
 XX
 XX 17-SEP-1993; 93US-0123859.
 XX
 XX (REGC) UNITV CALIFORNIA.
 XX
 XX Clare-Salzler MJ, Erlander MG, Kaufman DL, Toblin AJ;
 DR WPI; 1995-131360/17.
 XX
 XX New polypeptide fragments of glutamic acid decarboxylase - for
 PT diagnosis and treatment of autoimmune disease, esp. insulin
 PT dependent diabetes, also related nucleic acid, vectors,
 PT antibodies, hydridoma(s) etc.
 XX
 XX Claim 1; Page 76; 100pp; English.
 PS
 XX Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
 CC acid decarboxylase (GAD65) respectively, from which the GAD65
 CC fragments described in R72261-R72298 were derived. These fragments
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose
 CC and treat GAD-related autoimmune disorders, such as insulin
 CC dependant diabetes mellitus or stiff man disease.
 XX
 SQ Sequence 20 AA;
 OY 1 TLKYAIKTGHP 11
 DB 10 tlkyaiktghp 20
 RESULT 7 55.6%; Score 60; DB 16; Length 20;
 Y59551 ID Best Local Similarity 100.0%; Pred. No. 0.00044;
 XX ID T59551 standard; peptide: 20 AA.
 XX Y59551;
 AC
 XX
 XX 03-APR-2000 (first entry)
 DT
 XX
 DE GAD65 fragment, peptide #11.
 XX
 XX GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
 KW insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
 KW therapy.
 XX
 OS Homo sapiens.
 XX
 XX US5998366-A.
 PN
 XX
 XX 07-DEC-1999.
 PD
 XX
 XX 09-APR-1997; 97US-0827618.
 PF
 XX
 XX 07-JUN-1995; 95US-0485725.
 PR
 XX 21-SEP-1990; 90US-0586536.
 PR
 XX 18-JUN-1991; 91US-0716909.
 XX
 XX (REGC) UNITV CALIFORNIA.
 XX
 XX Toblin AJ, Kaufman DL, Erlander MG;
 PI
 XX
 DR WPI; 2000-095930/08.
 XX
 XX Ameliorating glutamic acid decarboxylase associated autoimmune
 PT disorders such as insulin dependent diabetes mellitus and Stiff man
 PT disease;
 XX
 PS Claim 1; Column 42; 61pp; English.

XX This sequence represents a fragment of the glutamic acid decarboxylase 65
 CC (GAD65) protein. The invention relates to a method of ameliorating GAD
 CC associated autoimmune disorders by administering a GAD65 peptide to the
 CC patient. The method can be used for ameliorating GAD associated
 CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
 CC and Stiff man disease. GAD65 can also be useful for screening drugs that
 CC alter GAD function, for generating monoclonal antibodies and in
 CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
 CC and the diagnosis is quite easy. It is also possible to obtain much
 CC larger quantities of polypeptide via recombinant techniques than are
 CC available from natural sources.
 XX
 SQ Sequence 20 AA;
 OY 1 TLKYAIKTGHP 11
 DB 10 tlkyaiktghp 20
 RESULT 8 55.6%; Score 60; DB 21; Length 20;
 W18846 ID Best Local Similarity 100.0%; Pred. No. 0.00044;
 XX ID W18846 standard; peptide: 20 AA.
 XX W18846;
 AC
 XX
 XX 05-JAN-1998 (first entry)
 DT
 XX
 DE 65 kD glutamic acid decarboxylase peptide fragment V.
 XX
 KW GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
 KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
 KW predisposition; autoimmune; tumour; rheumatoid arthritis;
 KW multiple sclerosis.
 XX
 OS Synthetic.
 XX
 XX DE19526561-A1.
 PN
 XX
 XX 23-JAN-1997.
 PD
 XX
 XX 20-JUL-1995; 95DE-1026561.
 PF
 XX
 XX 20-JUL-1995; 95DE-1026561.
 PR
 XX
 XX (BOE) BOEHRINGER MANNHEIM GMBH.
 XX
 XX Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;
 PI
 XX Pozzilli P, Stahl P;
 PI
 XX
 DR WPI; 1997-088254/09.
 XX
 XX Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
 PT involving intradermal admn. of auto:reactive substances
 PT
 XX
 XX Claim 11; Page 9; 12pp; German.
 PS
 XX W18842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predisposition to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 XX

SQ Sequence 20 AA:

Query Match 50.9%; Score 55; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.0033;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 PRYFNQLSTG 20
 |||||
 Db 1 pryfnglstg 10

RESULT 9

W01797
 ID W01797 standard; peptide; 20 AA.

AC W01797;

DT 15-OCT-1997 (first entry)

DE Human 65 kD glutamine decarboxylase peptide.

XX Human: glutamine decarboxylase; GAD; diagnosis; predisposition;
 KW tumour; immunological; disease; autoimmune; diabetes; reagent;
 KW determination; T cell; subpopulation; medication; treatment;
 KW prevention; production; antigen; immunogen; tolerogen; isolation;
 KW reinjection; inactivation.

OS Homo sapiens.

PN DE19525784-A1.

PD 16-JAN-1997.

PE 14-JUL-1995; 95DE-1025784.

PR 14-JUL-1995; 95DE-1025784.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Albert W, Boltard C, Endl J, Jung G, Schendel D;

PI Stahl P, Van Enderst P;

DR WPI: 1997-078452/08.

XX Glutamine decarboxylase peptide(s) - for diagnosis and therapy of
 PT diabetes, etc.

PS Claim 1; Page 12; 15pp; German.

XX The present peptide is a fragment of the human 65 kD glutamine
 CC decarboxylase (GAD), which can be used to diagnose, or diagnose a
 CC predisposition to, a tumour or immunological disease, preferably an
 CC autoimmune disease, especially diabetes. It can also be used as a
 CC reagent to determine specific T cell subpopulations, in medicaments
 CC to treat or prevent immunological diseases, preferably autoimmune
 CC diseases, especially diabetes, to produce antigens, especially
 CC immunogens or tolerogens and to isolate specific T cell
 CC subpopulations, which can be used to produce antigens or for
 CC reinjection, optionally after inactivation.

XX Sequence 20 AA:

Query Match 50.9%; Score 55; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.0033;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 PRYFNQLSTG 20
 |||||
 Db 1 pryfnglstg 10

RESULT 10

R76657
 ID R76657 standard; peptide; 14 AA.

AC R76657;

DT 05-MAR-1996 (first entry)

DE Peptide derived from human glutamic acid decarboxylase 16.

XX diabetes; T-cell subpopulation; detection; antigen production;
 KW diagnosis; autoimmune disease.

OS Homo sapiens.

PN DE4418091-A1.

PD 27-JUL-1995.

PE 24-MAY-1994; 94DE-4418091.

PR 04-FEB-1994; 94DE-4403522.

PR 20-JAN-1994; 94DE-4401629.

PA (ENDL) ENDL J.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Albert W, Dormair K, Endl J, Jung G, Mehl E;

PI Stahl P, Schendel D;

DR WPI: 1995-264505/35.

PS Claim 1; Fig 2; 21pp; German.

XX R76642-62 are derived from human glutamic acid decarboxylase and
 CC specifically react with T-cell subpopulations isolated from recently
 CC diagnosed Type-I diabetes. Pharmaceutical compns. contg. these
 CC peptides and those shown in R7571-72, are useful for the diagnosis of
 CC a disease or predispositions of immune system diseases, tumours, and
 CC autoimmune diseases, including diabetes. The peptides are able to detect
 CC specific T-cell subpopulations that are then used for antigen prodn.,
 CC e.g. by reinjection.

SQ Sequence 14 AA:

Query Match 44.4%; Score 48; DB 16; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 RYFNQLSTG 20
 |||||
 Db 1 ryinglstg 9

RESULT 11

W18865
 ID W18865 standard; peptide; 14 AA.

AC W18865;

DT 05-JAN-1998 (first entry)

DE 65 kD Glutamic acid decarboxylase peptide fragment 16.

XX GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
 KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
 KW predisposition; autoimmune; tumour; rheumatoid arthritis;
 KW multiple sclerosis.

```

XX OS Synthetic.
XX PN DE19526561-A1.
XX PD 23-JUN-1997.
XX PF 20-JUL-1995; 95DE-1026561.
XX PR 20-JUL-1995; 95DE-1026561.
XX PA (BOEF) BOEHRINGER MANNHEIM GMBH.
XX PI Donie F, Endl J, Ganz M, Jung G, Kientsch-engel R;
XX PI Pozzilli P, Stahl P;
XX DR WPI; 1997-088254/09.
XX PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
XX PT Involving Intra-dermal admn. of auto-reactive substances
XX PS Claim 11; Fig 2; 12pp; German.
XX CC W18842-10 are peptide fragments of the 65 kd human glutamic acid
XX CC decarboxylase (GAD). The fragments are autoreactive substances used for
XX CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
XX CC determined by using a claimed method for diagnosis of cell-mediated
XX CC diseases or a predisposition to cell-mediated diseases, which is effected
XX CC by administering an autoreactive substance intradermally and establishing
XX CC the diagnosis on the basis of the occurrence or lack of a positive
XX CC reaction at the site of administration. The method is used for diagnosis
XX CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
XX CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
XX SQ Sequence 14 AA;

Query Match 44.4%; Score 48; DB 18; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 RYFNQSTG 20
DB 1 rYfnqstg 9

RESULT 12
W12405
ID W12405 standard; peptide; 15 AA.
XX AC W12405;
XX AC 08-OCT-1997 (first entry)
XX DE GAD65 residues 177-191.
XX DE 65 kd glutamic acid decarboxylase; human; GAD65; enzyme; pancreas;
XX KW neuron; central nervous system; type I diabetes; autoimmune response;
XX KW T cell; therapy.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 1..15
XX FT /note- "optionally substituted, providing at least 7
XX FT residues remain wild type"
XX PN WO9700891-A1.
XX PD 09-JAN-1997.
XX PF 24-JUN-1996; 96WO-US10790.
XX

```

```

PR 23-JUN-1995; 95US-0494624.
XX PA (KENN-) KENNEDY INST RHEUMATOLOGY.
XX PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
XX PA (SBAR-) ST BARTHOLOMEW'S HOSPITAL CENT CLINICAL.
XX PI Conlon PJ, Gaur A, Leslie RDG, Ling N, Londel M;
XX PI WPI; 1997-087322/08.
XX PT New human glutamic acid decarboxylase peptide(s) - used for
XX PT treatment, diagnosis and determining predisposition to diabetes and
XX PT for ameliorating auto-immune responses.
XX PS Claim 23; Page -; 28pp; English.
XX CC W12403-W12413 represent fragments and analogues of the 65 kd isoform
XX CC of human glutamic acid decarboxylase (GAD65) (see W12402 for full length
XX CC wild type protein). GAD is an enzyme expressed in the beta cells of the
XX CC pancreas, and in neurons of the central nervous system. There are two
XX CC isoforms of GAD, a 67 kd isoform, and GAD65. Immunodominant regions of
XX CC GAD65 have been identified in Type I diabetic patients. These GAD65
XX CC fragments, and analogues, are used in the methods of the invention. The
XX CC methods are for detecting or treating diabetes or a predisposition to
XX CC diabetes. The peptides can also be used for ameliorating an autoimmune
XX CC response in a patient. Alteration of the native peptides with selective
XX CC changes of crucial residues can induce unresponsiveness or change the
XX CC responsiveness of antigen-specific autoreactive T cells. The peptide
XX CC analogues compete for binding to MHC and do not cause proliferation of
XX CC the corresponding native peptide-specific T cells.
XX SQ Sequence 15 AA;

Query Match 44.4%; Score 48; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 RYFNQSTG 20
DB 1 rYfnqstg 9

RESULT 13
Y36566
ID Y36566 standard; Protein; 24 AA.
XX AC Y36566;
XX DT 17-SEP-1999 (first entry)
XX DE Fragment of human secreted protein encoded by gene 39.
XX KW Human; secreted protein; cancer; tumour; developmental abnormality;
XX KW foetal deficiency; blood disorder; immune system disorder; inflammation;
XX KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
XX KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
XX KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
XX KW digestive disorder; endocrine disorder; infection; AIDS.
XX OS Homo sapiens.
XX PN WO9931117-A1.
XX PD 24-JUN-1999.
XX PF 17-DEC-1998; 98WO-US27059.
XX PF 19-DEC-1997; 97US-0068369.
XX PF 18-DEC-1997; 97US-0068006.
XX PF 18-DEC-1997; 97US-0068007.
XX PF 18-DEC-1997; 97US-0068008.
XX PF 18-DEC-1997; 97US-0068053.
XX

```


DR WPI: 1994-082825/10.
 XX Novel immunomodulatory peptide(s) and nucleic acids - useful for
 PT treatment of auto-immune diseases, transplant rejection and for
 PT vaccination
 PS Disclosure: Page 92: 139pp; English.
 XX
 CC The sequences given in R49291-505 and R46981-7038 represent peptide
 CC fragments of naturally-occurring immunomodulatory proteins. These
 CC fragments are between 10-30 residues in length and bind to a human
 CC major histocompatibility complex (MHC) class II allotype. These
 CC peptides may be used for therapy of autoimmune diseases, such as
 CC type I diabetes, rheumatoid arthritis and multiple sclerosis, and to
 CC reduce transplant rejection. They may also be used for vaccination
 CC providing an exclusively T-cell-mediated response, which can be
 CC class I or class-II based, or both, depending on the length and
 CC character of the immunogenic peptides.
 CC
 CC Sequence 24 AA:
 SQ
 Query Match 29.6%; Score 32; DB 15; Length 24;
 Best Local Similarity 60.0%; Pred. No. 43; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 4;
 QY 7 RTGHPRYFNQ 16
 Db 1 k1ghpdt1ng 10
 RESULT 16
 Y91936
 ID Y91936 standard; Peptide; 20 AA.
 AC Y91936:
 XX 19-JUL-2000 (first entry)
 DT
 DE HsRec2 peptide residues 153-172.
 XX
 KW HsRec2; serine kinase; homologous pairing; strand transfer; RAD51;
 KW phosphorylation; cyclin E; p53; cell cycle; substrate.
 XX
 OS Synthetic.
 XX
 PN WO200017329-A1.
 PD 30-MAR-2000.
 XX
 PF 17-SEP-1999; 99WO-US21642.
 XX
 PR 21-SEP-1998; 98US-0157603.
 XX
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 PA (CORR.) CORNELL RES FOUND INC.
 PA (KIME-) KIMERAGEN INC.
 PI
 PI Havre PA, Rice MC, Holloman WK, Kmiec EB;
 DR WPI: 2000-283562/24.
 XX
 XX Phosphorylating a serine-containing substrate by incubating it with
 PT adenosine triphosphate and Rec2 kinase and measuring the level of
 PT phosphorylation, useful for discovering specific antagonists or
 PT agonists of Rec2
 XX
 PS Example 3: Page 14; 41pp; English.
 XX The present sequence is hsrRec2 peptide, residues 153-172, a substrate
 CC which was not phosphorylated by hsrRec2, a human serine kinase. HsrRec2
 CC is in the same supergene family as the mammalian protein having
 CC homologous pairing and strand transfer activities, RAD51 and was

CC isolated because of its homology to the homologous pairing and strand
 CC transfer protein of *Ustilago maydis*. In particular, hsrRec2
 CC phosphorylates several proteins that control the cell cycle, especially
 CC cyclin E and p53. The invention permits the phosphorylation of cell
 CC cycle control proteins at sites that are physiologically relevant. The
 CC invention can be practiced with either murine or human Rec2 or a murine
 CC or chimeric of these proteins. In particular the murine has the sequence
 CC of a Rec2 kinase containing other than a Tyr at position 163. The
 CC invention comprises a method of phosphorylating a serine-containing
 CC substrate comprising incubating the substrate with ATP (adenosine
 CC triphosphate) and Rec2 kinase or a mammalian Rec2 and measuring the level
 CC of phosphorylation. The method is useful for discovering compounds which
 CC are specific antagonists or agonists of Rec2.
 CC
 CC Sequence 20 AA:
 SQ
 Query Match 28.7%; Score 31; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 52; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0;
 QY 11 PRYFN 15
 Db 9 prfyn 13
 RESULT 17
 Y57311
 ID Y57311 standard; peptide; 11 AA.
 XX
 AC Y57311:
 XX
 DT 05-JUN-2000 (first entry)
 DE P. aeruginosa ceramidase partial fragment C-86.
 XX
 KW Ceramidase; ceramide; sphingosine; fatty acid; atopic dermatitis;
 KW lipid engineering; anti-dermatitis.
 XX
 OS Pseudomonas aeruginosa.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 3 /note="unspecified"
 XX
 XX EP980912-A1.
 XX
 PN 23-FEB-2000.
 PD 20-AUG-1999; 99EP-0116154.
 XX
 PF 20-AUG-1998; 98JP-0234769.
 XX
 PR (TAKI) TAKARA SHUZO CO LTD.
 PA
 XX OKino N, Ito M;
 PI
 PI WPI: 2000-258590/23.
 DR
 XX New Pseudomonas aeruginosa alkaline ceramidase gene, useful for
 PT diagnosis, treatment and prevention of atopic dermatitis -
 PT
 XX Example 1: Page 21; 32pp; English.
 PS
 CC The invention provides an alkaline ceramidase from *P. aeruginosa*.
 CC Ceramidase can be produced by standard recombinant methodology.
 CC Ceramidase converts ceramide to sphingosine and a fatty acid, causes, or
 CC exacerbates, atopic dermatitis. The ceramidase gene, polypeptide, derived
 CC oligonucleotides and ceramidase-specific antibodies are used to prevent
 CC and/or treat atopic dermatitis. The gene and the polypeptide can be used
 CC as reagents for lipid engineering and for analyzing the structure and
 CC activity of ceramides. Sequences Y57308-313 represent partial fragments
 CC of the ceramidase protein, used for designing primers for amplifying the

CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of
 CC activity, allowing selection of agents for particular applications. Since
 CC these random copolymers induce total lysis of bacterial cell walls,
 CC resistance to them is unlikely to develop.
 XX
 SQ Sequence 16 AA;

Query Match 27.8%; Score 30; DB 18; Length 16;
 Best Local Similarity 50.0%; Pred. No. 59;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 10 HPRYFNQL 17
 |||| :
 Db 7 hprylkkl 14

RESULT 20
 W35210
 ID W35210 standard; peptide: 16 AA.

AC W35210;

DT 14-APR-1998 (first entry)

DE Diastereomer peptide 71.

KW Diastereomer peptide; infection; therapy; excitatory neurotoxin;

KW Honey bee venom; pardaxin; cytolytic activity; cancer;

KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
 KW agricultural pesticide; cell wall lysis.

OS Synthetic.

XX
 FH Key Location/Qualifiers

FT Misc-difference 2 /note= "D-form residue"

FT Misc-difference 3 /note= "D-form residue"

FT Misc-difference 4 /note= "D-form residue"

FT Misc-difference 7 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 10 /note= "D-form residue"

FT Misc-difference 11 /note= "D-form residue"

FT Misc-difference 13 /note= "D-form residue"

FT Misc-difference 14 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

FT Misc-difference 16 /note= "D-form residue"

XX Example 5; Page 47; 80pp; English.

PS This sequence represents a diastereomer peptide of the
 XX invention. The peptides of the invention have: (a) cytolytic activity on
 CC pathogenic cells (pathogens and malignant cells not naturally present in
 CC the body); but (b) no haemolytic activity, or such activity only at a
 CC concentration significantly higher than that at which they lyse
 CC pathogens. The peptides, their complexes and mixtures are used to treat
 CC infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
 CC or cancer, in human and veterinary medicine. Also, they can be used as
 CC preservatives for food, cosmetics and agricultural produce, or as
 CC agricultural pesticides. The absence of haemolytic activity (associated
 CC with disturbance of alpha-helical structures) means that the peptides
 CC have few if any toxic effects, and those that include D-aa will have
 CC increased resistance to proteolytic degradation. Non-haemolytic,
 CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of
 CC activity, allowing selection of agents for particular applications. Since
 CC these random copolymers induce total lysis of bacterial cell walls,
 CC resistance to them is unlikely to develop.

XX
 SQ Sequence 16 AA;

Query Match 27.8%; Score 30; DB 18; Length 16;
 Best Local Similarity 50.0%; Pred. No. 59;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 10 HPRYFNQL 17
 |||| :
 Db 7 hprylkkl 14

RESULT 21

W82934
 ID W82934 standard; peptide: 16 AA.

AC W82934;

DT 19-MAY-1999 (first entry)

DE Antipathogenic peptide.

KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;

KW cancer; infection; disinfectant; contact lens wetting solution;

KW preservative; pesticide; fungicide; bactericide.

OS Synthetic.

XX
 PN W09837090-A1.

PD 27-AUG-1998.

PF 19-FEB-1998; 98WO-IL00081.

PR 20-FEB-1997; 97WO-IL00066.

PA (YEDA) YEDA RES & DEV CO LTD.

PI Oren Z, Shai Y;

XX
 DR WPI: 1998-594464/50.

XX New non-haemolytic cytolytic agent useful in treating cancer or

XX infections - is a peptide comprising a moiety which disrupts the

XX continuity of an alpha-helical structure

XX
 PS Example 5; Page 46; 126pp; English.

XX The present peptide is used to produce the agents of the invention. The

XX specification describes a non-haemolytic, cytolytic agent, which is a

XX peptide, a complex of bundled peptides, a mixture of peptides or a random

XX peptide copolymer. The agent has a selective cytolytic activity on

XX

CC pathogenic cells. The agent is selected from a cyclic derivative of a
 CC peptide which has a net positive charge greater than 1, comprises L-amino
 CC acid residues and/or D-amino acid residues and comprises an alpha-helix
 CC breaker moiety, or a peptide (or cyclic derivative of this) which
 CC (comprises L-amino acid residues and D-amino acid residues, has a net
 CC positive charge greater than 1 and has an amino acid sequence such that
 CC a corresponding amino acid sequence comprising only L-amino acid residues
 CC is not found in nature. The cytolytic agents may be used for treatment of
 CC cancer or for treatment of several diseases caused by pathogens,
 CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
 CC They may be used in both human and veterinary medicine. They may also be
 CC used as disinfectants for destruction of microorganisms, i.e. in
 CC solutions for wetting contact lenses, as preservatives, e.g., in the
 CC cosmetic and food industries, as pesticides (e.g. fungicides or
 CC bactericides) or for preservation of agricultural products.

XX Sequence 16 AA;

Query Match 27.8%; Score 30; DB 19; Length 16;
 Best Local Similarity 50.0%; Pred. No. 59;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 10 HPRYFNOL 17
 DB 7 hpryikkl 14

RESULT 22

W82926 W82926 standard; peptide; 16 AA.

AC W82926;

DT 19-MAY-1999 (first entry)

DE Antipathogenic peptide.

XX Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;

KW cancer; infection; disinfectant; contact lens wetting solution;

XX preservative; pesticide; fungicide; bactericide.

OS Synthetic.

PN W09837090-A1.

PD 27-AUG-1998.

PF 19-FEB-1998; 98WO-1100081.

PR 20-FEB-1997; 97WO-1100066.

PA (YEDA) YEDA RES & DEV CO LTD.

PI Oren Z, Shai Y;

DR WPI; 1998-594464/50.

XX New non-haemolytic cytolytic agent useful in treating cancer or

PT infections - is a peptide comprising a moiety which disrupts the

XX continuity of an alpha-helical structure

PS Example 5; Page 46; 126pp; English.

CC The present peptide is used to produce the agents of the invention. The
 CC specification describes a non-haemolytic, cytolytic agent, which is a
 CC peptide, a complex of bundled peptides, a mixture of peptides or a random
 CC peptide copolymer. The agent has a selective cytolytic activity on
 CC pathogenic cells. The agent is selected from a cyclic derivative of a
 CC peptide which has a net positive charge greater than 1, comprises L-amino
 CC acid residues and/or D-amino acid residues and comprises an alpha-helix
 CC breaker moiety, or a peptide (or cyclic derivative of this) which
 CC (comprises L-amino acid residues and D-amino acid residues, has a net

CC positive charge greater than 1 and has an amino acid sequence such that
 CC a corresponding amino acid sequence comprising only L-amino acid residues
 CC is not found in nature. The cytolytic agents may be used for treatment of
 CC cancer or for treatment of several diseases caused by pathogens,
 CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
 CC They may be used in both human and veterinary medicine. They may also be
 CC used as disinfectants for destruction of microorganisms, i.e. in
 CC solutions for wetting contact lenses, as preservatives, e.g., in the
 CC cosmetic and food industries, as pesticides (e.g. fungicides or
 CC bactericides) or for preservation of agricultural products.

XX Sequence 16 AA;

Query Match 27.8%; Score 30; DB 19; Length 16;
 Best Local Similarity 50.0%; Pred. No. 59;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 10 HPRYFNOL 17
 DB 7 hpryikkl 14

RESULT 23

W80790 W80790 standard; peptide; 17 AA.

AC W80790;

DT 07-DEC-1998 (first entry)

DE Peptide 6 from D. discoideum DPI protein.

XX Dictyostelium; DPI; discoidin II; human DPI; hDPI; cyclophilin; Ddcyp2;

KW bisphosphonate binding protein; calcium metabolism; cyclosporine;

XX osteoporosis; hypercalcaemia; bone metabolism; bone metastases.

OS Dictyostelium discoideum.

XX Key Location/Qualifiers

FT Misc-difference 1 /note="unknown"

PN W09836064-A1.

PD 20-AUG-1998.

PF 13-FEB-1998; 98WO-US02709.

PR 14-FEB-1997; 97US-0039738.

PA (PROC) PROCTER & GAMBLE CO.

XX (UYSH-) UNIV SHEFFIELD MED SCHOOL.

PI Cook JS, Ebelino FH, Ibbotson KJ, Ji X, Rogers MJ;

PT Russell RGR, Watts DJ, Xiong XJ;

DR WPI; 1998-506311/43.

XX Bisphosphonate binding protein - used to treat calcium metabolism

PT disorders, including bone metabolism, hypercalcaemia, bone

XX metastases, and osteoporosis

PS Example 1; Page 28; 98pp; English.

CC Sequences shown in W80785 to W80793 represent peptides from DPI produced
 CC by endoprotease digestion, used for peptide sequencing. The invention
 CC provides sequences encoding Dictyostelium DPI, human DPI (hDPI) and
 CC Dictyostelium cyclophilin (Ddcyp2) which are bisphosphonate binding
 CC proteins. The invention also provides methods for purifying and producing
 CC such bisphosphonate binding proteins. The bisphosphonate binding protein,
 CC or an antibody which binds to the binding protein can be used in the
 CC diagnosis of calcium metabolism disorders. They can also be used to treat

CC calcium metabolism disorders, where the treatment is for the regulation
 CC of bone metabolism, hypercalcaemia, bone metastases, and osteoporosis,
 CC especially by regulation of bone metabolism via interaction with
 CC cyclosporine.

XX Sequence 17 AA;

Query Match 27.8%; Score 30; DB 19; Length 17;
 Best Local Similarity 55.6%; Pred. No. 64;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 10 HPRFNQLS 18
 ||| | : |
 Db 7 hptynh15 15

RESULT 24

Y01008 Y01008 standard; peptide; 9 AA.

XX Y01008;

DT 02-JUN-1999 (first entry)

DE Bacterial immunogenic peptide clone 18c.

KW Immunogenic peptide; epitope unit; immune response; in vivo protection;
 KW pathogenic microorganism; microorganism infection; diagnosis; therapy.

OS Synthetic.

PN W09911660-A2.

PD 11-MAR-1999.

PF 03-SEP-1998; 98WO-EP05575.

PR 04-SEP-1997; 97US-0057906.

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX (INSP) INST PASTEUR.

PI Cortese R, Felici F, Kraehenbuhl J, Phalipon A;

PI Sansonetti P;

DR WPI; 1999-205131/17.

PT New immunogenic polypeptide which induces an immune response in vivo
 PT - useful for inducing a stronger protective immune response against
 PT pathogenic microorganisms

XX Disclosure; Fig 5; 64pp; English.

CC This sequence represents an immunogenic polypeptide of the invention, and
 CC is capable of inducing an immune response in vivo against a pathogenic
 CC microorganism. The peptides can be used for pharmaceutical compositions
 CC for in vivo protection against pathogenic microorganisms. Antibodies
 CC recognising the peptides are useful for diagnosing microorganism
 CC infections. The peptides induce a stronger protective immune response to
 CC pathogenic microorganisms compared to prior art peptides. Preparation is
 CC cheaper and less time is consumed than in preparation of polysaccharide
 CC or carbohydrate antigens.

XX Sequence 9 AA;

Query Match 26.9%; Score 29; DB 20; Length 9;

Best Local Similarity 57.1%; Pred. No. 2.1e+05;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 11 PRYFNQL 17
 | : ||| :

Db 2 phffnql 8

RESULT 25

ID W92558 standard; peptide; 15 AA.

XX W92558;

DT 26-APR-1999 (first entry)

DE Mouse beta-actin array peptide substrate #9.

KW Peptide substrate; CCR; eukaryotic type II chaperonin complex; cyclin;
 KW binding agent; substrate-binding site; SBS; substrate folding; actin;
 KW tubulin; treatment; cancer; anticancer drug; viral infection; screening;
 KW reduced toxicity; mouse.

OS Mus sp.

PN W09853322-A1.

PD 26-NOV-1998.

PF 22-MAY-1998; 98WO-GB01485.

PR 23-MAY-1997; 97GB-0010762.

PA (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.

XX Hynes G, Llou AK, Wallison K;

DR WPI; 1999-070162/06.

PT Identifying specific binding agents for substrate binding site in
 PT CCR chaperonin complex - also new peptide binding agents and their
 PT mimetics, and peptides containing a specific CCR binding site, used
 PT for treating cancer

XX Disclosure; Page 87; 97pp; English.

CC This invention describes a method which uses the CCR (eukaryotic type II
 CC chaperonin) complex or part of it, for identifying a binding agent that
 CC can occupy a substrate-binding site (SBS) on the CCR complex. By binding
 CC to the CCR complex, the binding agents block an SBS so that biological
 CC activity of the CCR complex is affected, particularly its ability to fold
 CC substrates such as actin, tubulin and cyclin. The binding agents are
 CC useful for treatment of cancer, particularly when used in combination
 CC with an anticancer drug, or viral infections. Nucleic acid fragments are
 CC between the CCR complex and a protein that is to be folded. The binding
 CC agents may target cells that are actively synthesising tubulin etc.
 CC (unlike known microtubule-stabilising agents that affect all cells), so
 CC should have reduced toxicity for normal cells. W92550-W92628 are
 CC peptide substrates used in the method of the invention.

XX Sequence 15 AA;

Query Match 26.9%; Score 29; DB 20; Length 15;

Best Local Similarity 66.7%; Pred. No. 82;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 TLKTAIKTG 9
 ||||| : |
 Db 6 tlkylpleng 14

Search completed: February 5, 2001, 10:47:33
 Job time: 627 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:49:30 ; Search time 58.67 Seconds
(without alignments)
23.147 Million cell updates/sec

Title: US-08-981-824-5

Perfect score: 109
Sequence: 1 PRYFNQSTGLDVGGLADW 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 5001

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	28.4	25	2	gene HEXA protein
2	28	25.7	15	2	phosphoribulokinase
3	28	25.7	25	2	alpha-2u-globulin
4	27	24.8	13	2	urinary tract ston
5	27	24.8	16	2	protein kinase C 1
6	26	23.9	21	2	gene CTLA-1 protei
7	26	23.9	21	2	NF1/CRAF-binding t
8	25	22.9	13	2	cytochrome P450 CM
9	25	22.9	16	2	major outer membra
10	25	22.9	19	2	homeobox 3 protein
11	25	22.9	19	2	homeobox 4 protein
12	25	22.9	20	2	comartin 7-monooxy
13	25	22.9	20	2	kinase-related tra
14	25	22.9	23	2	photosystem I 8.0K
15	25	22.9	25	2	ubiquitin thiolest
16	25	22.9	25	2	aryldialkylphospa
17	24	22.0	12	2	gene HEXA protein
18	24	22.0	15	2	allatostatin - tob
19	24	22.0	16	2	multicatalytic end
20	24	22.0	18	2	TCR delta chain V-
21	24	22.0	20	2	multicatalytic end
22	24	22.0	23	2	exo-poly-alpha-gal
23	24	22.0	25	2	amide (EC 3.5.1.
24	24	22.0	17	2	hypothetical prote
25	23	21.1	17	2	Ig heavy chain CRD
26	23	21.1	19	2	UDPglucose--glycos
27	23	21.1	19	2	photosystem I 8.0K
28	23	21.1	20	2	39k major outer me
29	23	21.1	23	2	heterodisulfide re

30	23	21.1	24	2	S34405
31	23	21.1	24	2	S66561
32	23	21.1	24	2	A41037
33	22	20.2	12	2	S34447
34	22	20.2	15	2	A48372
35	22	20.2	15	2	A26212
36	22	20.2	16	2	S10809
37	22	20.2	16	2	S09700
38	22	20.2	16	2	A26393
39	22	20.2	17	2	A58946
40	22	20.2	17	2	S19614
41	22	20.2	19	2	JC2057
42	22	20.2	21	2	S28436
43	22	20.2	24	2	PC2305
44	22	20.2	24	2	S59493
45	22	20.2	24	2	S52034

ALIGNMENTS

RESULT 1
164828
gene HEXA protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C:Accession: 164828
R:Boles, D.J.; Proia, R.L.
Am. J. Hum. Genet. 56: 716-724, 1995
A:Title: The molecular basis of HEXA mRNA deficiency caused by the most common Tay-Sa
A:Reference number: 151862; MUID:95193801
A:Accession: 164828
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-25 <RES>
A:Cross-references: GB:S76982; NID:9912780; PIDN:AD14242.1; PID:94261942
C:Genetics:
A:Gene: GDB:HEXA
A:Cross-references: GDB:120040; OMIM:272800
A:Map position: 15q23-15q24
C:Superfamily: beta-hexosaminidase

Query Match 28.4%, Score 31; DB 2; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OR 1 PRYFNQSTGLD 12
DB 1 PWYINRISYGPD 12
RESULT 2
PS0276
phosphoribulokinase (EC 2.7.1.19) - rice (strain Nihonbare) (fragment)
C:Species: Oryza sativa (rice)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Mar-1995
C:Accession: PS0276
R:Tangita, A.
submitted to JIPID, April 1993
A:Reference number: PS0206
A:Accession: PS0276
A:Molecule type: protein
A:Residues: 1-15 <TSD>
A:Experimental source: leaf, chlorophyll, stem
C:Keywords: phosphotransferase

Query Match 25.7%, Score 28; DB 2; Length 15;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OR 13 MWGLAAD 19

DB 7 VIGLAAD 13

RESULT 3

S70348
alpha-2u-globulin - rat (fragment)
N:Alternate names: alpha-2-uglobulin; major urinary protein
C:Species: Rattus norvegicus (Norway rat)
C>Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C:Accession: S70348
R:Bayard, C.; Holmquist, L.; Vesterberg, O.
Biochim. Biophys. Acta 1290, 129-134, 1996
A:Title: Purification and identification of allergenic alpha(2u)-globulin species of rat
A:Reference number: S70347; MUID:96254071
A:Accession: S70348
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <BAV>
C:Superfamily: lipocalin; lipocalin homology

Query Match 25.7%; Score 28; DB 2; Length 25;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 11 LDWVGLADW 20
||| : ||
Db 10 LDVAKLNGDW 19

RESULT 4

H56046
urinary tract stone matrix protein 10, 42k - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 12-Apr-1995
C:Accession: H56046
R:Binette, J.P.; Binette, M.B.; Gawinowicz, M.A.; Kendrick, N.
Submitted to the Protein Sequence Database, February 1995
A:Description: Isolation, characterization and sequence of stone proteins.
A:Reference number: A56046
A:Accession: H56046
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-13 <BIN>

Query Match 24.8%; Score 27; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 YFNOLSTGLD 12
||| : |
Db 4 YFNDLAETD 13

RESULT 5

S10807
protein kinase C inhibitor KcIP-1 isoform d - sheep (fragment)
N:Alternate names: probable lipocortin
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 19-Mar-1997 #sequence_revision 13-Sep-1998 #text_change 13-Sep-1998
C:Accession: S10807
R:Tooker, A.; Ellis, C.A.; Sellers, L.A.; Altken, A.
Eur. J. Biochem. 191, 421-429, 1990
A:Title: Protein kinase C inhibitor proteins. Purification from sheep brain and sequence
A:Reference number: S10804; MUID:90345949
A:Accession: S10807
A:Molecule type: protein
A:Residues: 1-16 <TK>
A:Experimental source: brain
C:Superfamily: 14-3-3 protein

Query Match 24.8%; Score 27; DB 2; Length 16;
Best Local Similarity 45.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYFNOLSTGLD 12
||| : |||
Db 6 RYLAERATGND 16

RESULT 6

I49414
gene CTLA-1 protein - western wild mouse (fragment)
C:Species: Mus spretus (western wild mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1999
C:Accession: I49414
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, Mamm. Genome 5, 349-355, 1994
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A:Reference number: I4934; MUID:94319082
A:Accession: I49414
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-21 <RES>
A:Cross-references: EMBL:U05708; NID:q497037; PIDN:AA60471.1; PID:q497038
C:Genetics:
A:gene: CTLA-1
C:Superfamily: trypsin; trypsin homology

Query Match 23.9%; Score 26; DB 2; Length 21;
Best Local Similarity 35.7%; Pred. No. 6.4e+02;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 PRYFNOLSTGLDW 14
||| : |||
Db 1 PRAFTKVSFLSWI 14

RESULT 7

S69502
NFI/CAAT-binding transcription factor 5 (CTF5) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C:Accession: S69502
R:Wenzelides, S.; Altman, H.; Wendler, W.; Winacker, E.L.
Nucleic Acids Res. 24, 2416-2421, 1996
A:Title: CTF5 - a new transcriptional activator of the NFI/CTF family.
A:Reference number: S69502; MUID:96279745
A:Accession: S69502
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-21 <WEN>
A:Cross-references: EMBL:X92857
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

Query Match 23.9%; Score 26; DB 2; Length 21;
Best Local Similarity 62.5%; Pred. No. 6.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 12 DMWGLAAD 19
||| : |||
Db 1 DLVSLACD 8

RESULT 8

S36874
cytochrome P450 CMB (EC 1.14.-.-) - crab-eating macaque (fragment)
C:Species: Macaca fascicularis (crab-eating macaque)
C>Date: 09-Jun-1994 #sequence_revision 07-Nov-1997 #text_change 05-Mar-1999
C:Accession: S36874
R:Ohmori, S.; Horie, T.; Guengerich, F.P.; Kluchl, M.; Kitada, M.

Arch. Biochem. Biophys. 305, 405-413, 1993
 A:Title: Purification and characterization of two forms of hepatic microsomal cytochrome
 A:Reference number: S36874; MUID:93384294
 A:Accession: S36874
 A:Molecule type: protein
 A:Residues: 1-13 <OHM>
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
 C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; liver;

Query Match 22.9%; Score 25; DB 2; Length 13;
 Best Local Similarity 50.0%; Pred. No. 5.6e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 7 LSTGLDNYGLAA 18
 |||||
 Db 2 LAGGLLVALLA 13

RESULT 9
 A49226

major outer membrane protein - Haemophilus somnus (fragment)
 C:Species: Haemophilus somnus
 C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994

R:Tagawa, Y.; Ishikawa, H.; Yuasa, N.

Infect. Immun. 61, 91-96, 1993

A:Title: Purification and partial characterization of the major outer membrane protein C
 A:Reference number: A49226; MUID:93114910

A:Contents: 8025

A:Accession: A49226

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-16 <TAG>
 A:Note: sequence extracted from NCBI backbone (NCBIF:121595)

Query Match 22.9%; Score 25; DB 2; Length 16;
 Best Local Similarity 41.7%; Pred. No. 7e+02;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 4 FNOLSTGLDNG 15
 |||||
 Db 4 YNNGTKYDVGC 15

RESULT 10
 JC2058

homeobox 3 protein - common tobacco (fragment)

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Aug-1994

C:Accession: JC2058

R:Feng, X.H.; Kung, S.D.

Biochem. Biophys. Res. Commun. 198, 1012-1019, 1994

A:Title: Identification of differentially expressed members of tobacco homeobox families
 A:Reference number: JC2057; MUID:94161708

A:Accession: JC2058

A:Molecule type: DNA

A:Residues: 1-19 <FEN>

Query Match 22.9%; Score 25; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STGLD 12
 |||||
 Db 10 STGLD 14

RESULT 11

JC2059
 homeobox 4 protein - common tobacco (fragment)

C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 23-Mar-1995

C:Accession: JC2059

R:Feng, X.H.; Kung, S.D.

Biochem. Biophys. Res. Commun. 198, 1012-1019, 1994

A:Title: Identification of differentially expressed members of tobacco homeobox fam11
 A:Reference number: JC2057; MUID:94161708

A:Accession: JC2059

A:Molecule type: DNA

A:Residues: 1-19 <FEN>

A:Experimental source: leaf

C:Genetics:

A:Gene: Hc4

C:Keywords: homeobox

Query Match 22.9%; Score 25; DB 2; Length 19;
 Best Local Similarity 70.0%; Pred. No. 8.5e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

OY 5 NQLS--TGID 12
 |||||
 Db 5 NRLSETGLD 14

RESULT 12

S21737
 coumarin 7-monooxygenase (EC 1.14.14.-) cytochrome P450 2A7 - western baboon (fragmen

N:Alternate names: coumarin 7-hydroxylase

C:Species: Papio papio (western baboon)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Mar-1999

C:Accession: S21737

R:Dalet-Beluche, I.; Boulenc, X.; Fabre, G.; Mauriel, P.; Bonfils, C.

Eur. J. Biochem. 204, 641-648, 1992

A:Title: Purification of two cytochrome P450 isozymes related to CYP2A and CYP3A gene

A:Reference number: S21737; MUID:92174920

A:Accession: S21737

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <DAL>

C:Genetics:

A:Gene: CYP2A7

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
 C:Keywords: heme; microsome; monooxygenase; oxidoreductase; transmembrane protein

Query Match 22.9%; Score 25; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 9e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 7 LSTGLDNYGLAA 18
 |||||
 Db 2 LAGGLLVALLA 13

RESULT 13

S00774
 kinase-related transforming protein (abl) (EC 2.7.1.-) type II - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 18-Jun-1999

C:Accession: S00774

R:Bernards, A.; Pasikind, M.; Baltimore, D.

Oncogene 2, 297-304, 1988

A:Title: Four murine c-abl mRNAs arise by usage of two transcriptional promoters and
 A:Reference number: S00771; MUID:88202920

A:Accession: S00774

A:Molecule type: DNA

A:Residues: 1-20 <BER>

A:Cross-references: EMBL:X07540; MID:949835; PIDN:CA30413.1; PID:949836

A:Note: the authors translated the codon GAC for residue 12 as His

C:Genetics:

A:Gene: abl

C:Superfamily: human protein-tyrosine kinase abl; protein kinase homology; SH2 homoio

C:Keywords: alternative splicing; ATP; phosphotransferase

Query Match 22.9%; Score 25; DB 2; Length 20;
Best Local Similarity 43.8%; Pred. No. 9e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 4 FNOLSTGLDMVGLAD 19
| | | | |
Db 4 FDLISDELDKLVLVD 19

RESULT 14

P00690
Photosystem I 8.0K G chain - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
C:Accession: P00690
R:Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Sugitara, M.
Plant. Physiol. 102, 1259-1267, 1993
A:Title: Molecular heterogeneity of photosystem I. psad, psae, psaf, psah and psal are
A:Reference number: P00667; MWID:94105345
A:Accession: P00690
A:Molecule type: protein
A:Residues: 1-23 <OBO>
C:Superfamily: photosystem I chain V
C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match 22.9%; Score 25; DB 2; Length 23;
Best Local Similarity 71.4%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 LSTGLDM 13
| | | | |
Db 10 LSTGLSL 16

RESULT 15

S17561
ubiquitin thiolesterase (EC 3.1.2.15) PGP9.5, brain - bovine (fragment)
N:Alternate names: 26K protein; gene product 9.5; ubiquitin carboxyl-terminal esterase
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 13-Sep-1998 #text_change 11-Jan-2000
C:Accession: S17561
R:Giambanco, I.; Bianchi, R.; Ceccarelli, P.; Pula, G.; Sorci, G.; Antonioi, S.; Bocchi,
FERS Lett. 290, 131-134, 1991
A:Title: Neuron-specific protein gene product 9.5 (PGP 9.5) is also expressed in glioma
A:Reference number: S17561; MWID:92008646
A:Accession: S17561
A:Molecule type: protein
A:Residues: 1-25 <GIA>
A:Experimental source: brain; glioma cells
C:Genetics:
A:Gene: 9.5
C:Function:

A:Description: involved in the hydrolysis of esters and amides at the C-terminal glycine
e during brain development in glial cells
C:Superfamily: human ubiquitin thiolesterase
C:Keywords: cytosol; thiolester hydrolase

Query Match 22.9%; Score 25; DB 2; Length 25;
Best Local Similarity 33.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 PXYENOLSTGLDMVG 15
| | | | |
Db 10 PEMLNKVLTRLGAVG 24

RESULT 16

P00088
Aryldialkylphosphatase (EC 3.1.8.1) - rat (fragment)
N:Alternate names: A-esterase; Aryltriphosphatase; Organophosphate hydrolase; Paraoxa
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997
C:Accession: P00088
R:Toungou, K.
submitted to JIPID, January 1996
A:Reference number: P00088
A:Accession: P00088
A:Molecule type: protein
A:Residues: 1-25 <TOU>
C:Keywords: phosphoric triester hydrolase

Query Match 22.9%; Score 25; DB 2; Length 25;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 10 GLDMGL 16
| | | | |
Db 5 GLTLVGL 11

RESULT 17
I64829
gene HEXA protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C:Accession: I64829
R:Boles, D.J.; Proia, R.L.
Am. J. Hum. Genet. 56, 716-724, 1995
A:Title: The molecular basis of HEXA mRNA deficiency caused by the most common Tay-Sa
A:Reference number: I51882; MWID:95193801
A:Accession: I64829
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-12 <RES>
A:Cross-references: GB:S76984; NID:9912781; PIDN:AA014243.1; PID:94261943
C:Genetics:
A:Gene: GDB:HEXA
A:Cross-references: GDB:120040; OMIM:272800
A:Map position: 15q23-15q24
C:Superfamily: beta-hexosaminidase

Query Match 22.0%; Score 24; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 PXYENOLS 8
| | | | |
Db 1 PWTYLRIS 8

RESULT 18
A61612
allatostatin - tobacco hornworm
C:Species: Manduca sexta (tobacco hornworm)
C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 11-Jul-1997
C:Accession: A61612
R:Kramer, S.J.; Toschl, A.; Miller, C.A.; Kataoka, H.; Quistad, G.B.; Li, J.P.; Carne
Proc. Natl. Acad. Sci. U.S.A. 88, 9458-9462, 1991
A:Title: Identification of an allatostatin from the tobacco hornworm Manduca sexta.
A:Reference number: A61612; MWID:92052112
A:Accession: A61612
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-15 <KRA>
C:Keywords: neuropeptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 15;

Best Local Similarity 66.7%; Pred. No. 9.6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 YFNOLS 8
|||:|
Db 8 YFNPI5 13

RESULT 19

PC3320 multicatalytic endopeptidase complex (EC 3.4.99.46), HC10-II chain - human (fragment)

N:Alternate names: proteasome
C:Species: Homo sapiens (man)
C:Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 26-May-2000
C:Accession: PC3320

R:Kristensen, P.; Johnsen, A.H.; Uerkvitz, M.; Tanaka, K.; Hendll, K.B.

Biochem. Biophys. Res. Commun. 205, 1785-1789, 1994

A:Title: Human proteasome subunits from 2-dimensional gels identified by partial sequence
A:Reference number: PC3315; MUID:95110324

A:Accession: PC3320

A:Molecule type: protein

A:Residues: 1-16 <KRI>

A:Experimental source: placenta

C:Comment: The proteasome consists of subunits of 21k-30k arranged in 4 stacked rings.

C:Superfamily: multicatalytic endopeptidase complex chain C9

C:Keywords: hydrolase; proteinase.

Query Match 22.0%; Score 24; DB 2; Length 16;
Best Local Similarity 41.7%; Pred. No. 1e+03;

Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 PRYFNOLSTGLD 12
| | | | |
Db 3 PYTEPVVIGLD 14

RESULT 20

G49037 Tcr delta chain V-D-J region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: G49037

R:Rezqueria, A.; Wilde, D.B.; McConnell, T.J.; Sturmhofel, K.; Valas, R.B.; Shevach, E.M.
Eur. J. Immunol. 22, 491-498, 1992

A:Title: Mouse autoreactive gamma/delta T cells. II. Molecular characterization of the T

A:Reference number: A49037; MUID:92164730

A:Accession: G49037

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-18 <EAO>

A:Cross-references: GB:S90657; NID:q246300; PIDN:AB21553.1; PID:q246301

A:Experimental source: dendritic epidermal T-cell lines

A>Note: sequence extracted from NCBI backbone (NCBIN:90657, NCBIIP:90668)

Query Match 22.0%; Score 24; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;

Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 10 GLDMGLAD 19
| | | | |
Db 7 GISRGLATD 16

RESULT 21

F42762 multicatalytic endopeptidase complex (EC 3.4.99.46) subunit 13 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 26-May-2000
C:Accession: F42762

R:Dick, L.R.; Moormaw, C.R.; Pramanik, B.C.; DeMartino, G.N.; Slaughter, C.A.
Biochemistry 31, 7347-7355, 1992

A:Title: Identification and localization of a cysteinyl residue critical for the try

A:Reference number: A42762; MUID:92378961

A:Accession: F42762

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <DIC>

A>Note: sequence extracted from NCBI backbone (NCBIIP:112175)

C:Superfamily: multicatalytic endopeptidase complex chain C9

C:Keywords: hydrolase

Query Match 22.0%; Score 24; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 10 GLDMGLAD 19
| | | | |
Db 3 GRNXVAVIAD 12

RESULT 22

A48968 exo-poly-alpha-galacturonosidase (EC 3.2.1.82) - Clostridium thermosaccharolyticum (f

N:Alternate names: exo-poly-alpha-galacturonate hydrolase

C:Species: Clostridium thermosaccharolyticum, Clostridium tartarivorum

C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 06-Dec-1996
C:Accession: A48968

R:Van Rijssel, M.; Gerwig, G.J.; Hansen, T.A.

Appl. Environ. Microbiol. 59, 828-836, 1993

A:Title: Isolation and characterization of an extracellular glycosylated protein comp

A:Reference number: A48968; MUID:93243739

A:Accession: A48968

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-23 <VAN>

A>Note: sequence extracted from NCBI backbone (NCBIIP:130462)

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 22.0%; Score 24; DB 2; Length 23;
Best Local Similarity 40.0%; Pred. No. 1.5e+03;

Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 ENQSTGLDM 13
| | | | |
Db 12 FNDIXTSNDL 21

RESULT 23

PC1105 amylase (EC 3.5.1.4) - Brevibacterium sp. (fragment)

C:Species: Brevibacterium sp.

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: PC1105

R:Soulier, F.; Levy-Schil, S.; Mayeux, J.F.; Petre, D.; Arnaud, A.; Crouzet, J.

Gene 116, 99-104, 1992

A:Title: Cloning and primary structure of the wide-spectrum amylase from Brevibacteri

A:Reference number: JC1174; MUID:92331957

A:Accession: PC1105

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-25 <SDU>

C:Keywords: hydrolase

Query Match 22.0%; Score 24; DB 2; Length 25;
Best Local Similarity 45.5%; Pred. No. 1.7e+03;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 7 LSTGLDMVGLA 17
| | | | |
Db 6 ISSSNDTVGVA 16

RESULT 24

S58129

hypothetical protein - moss (Ceratodon purpureus)

C:Species: Ceratodon purpureus

C:Date: 13-Jan-1996

R: Hughes, J.; Mittleman, F.

submitted to the EMBL Data Library, July 1995

A:Description: The moss Ceratodon purpureus contains and expresses a second, conventional

A:Accession: S58129

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-17 <HUG>

A:Cross-references: GB:U56698; EMBL:X89725; NID:g1314836; PID:g1322246

Query Match 21.1%; Score 23; DB 2; Length 17;

Best Local Similarity 38.5%; Pred. NO. 1.6e+03;

Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 4 FMQLSTGLDWGL 16

DB 4 FSTSRSLMIVGI 16

RESULT 25

PT0234

Ig heavy chain CDR3 region (clone 1-130) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993

R: Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and

A:Reference number: PT0222; MUID:91108337

A:Accession: PT0234

A:Molecule type: DNA

A:Residues: 1-17 <YAM>

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match 21.1%; Score 23; DB 2; Length 17;

Best Local Similarity 55.6%; Pred. NO. 1.6e+03;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 RYFNOLSTG 10

DB 8 RYFDWLGG 16

Search completed: February 5, 2001, 10:49:31
Job time: 744 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:56:01 ; Search time 32.57 Seconds
(without alignments)
19.831 Million cell updates/sec

Title: US-08-981-824-5

Sequence: 1 PRYFNQLSTGLDMVGLADW 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 1436

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	26.6	14	1	KPPI_SELMI
2	28	25.7	15	1	KPPI_SELMI
3	27	24.8	17	1	RANR_RANRU
4	26	23.9	23	1	LEO1_BIOGL
5	25	22.9	20	1	CPA7_PAPSP
6	25	22.9	25	1	UBL1_BOVIN
7	24	22.0	15	1	ALIS_MANSE
8	23	21.1	9	1	OXYF_SCYCA
9	23	21.1	19	1	PPBH_PSEAE
10	23	21.1	20	1	OMP1_ACTAC
11	23	21.1	24	1	KAD_BACLI
12	23	21.1	24	1	PGO_XENLA
13	23	21.1	25	1	CR21_LITSP
14	23	21.1	25	1	CR22_LITGI
15	23	21.1	25	1	CR23_LITCE
16	23	21.1	25	1	CR24_LITCE
17	23	21.1	25	1	CR25_LITGI
18	22	20.2	15	1	CBPB_PROAT
19	22	20.2	19	1	OXLA_OPHNA
20	22	20.2	20	1	FTBR_PACIE
21	22	20.2	21	1	SOD2_PICAB
22	22	20.2	24	1	ALU4_OLEEU
23	22	20.2	24	1	GA66_RANRU
24	21.5	19.7	18	1	NODD_RHTIO
25	21	19.3	15	1	CDN2_LITGI
26	21	19.3	18	1	CTIA_LITGI
27	21	19.3	18	1	CTIB_LITGI
28	21	19.3	18	1	CTIC_LITCI
29	21	19.3	18	1	CTID_LITCI
30	21	19.3	18	1	YPE4_LIACI
31	21	19.3	22	1	MYSB_CANFA
32	21	19.3	25	1	K6P2_THERH
33	20	18.3	19	1	IRBP_CAVPO

34	20	18.3	21	1	BOH1_BOMVA
35	20	18.3	22	1	ATP6_COTJA
36	20	18.3	22	1	SETB_SALTY
37	20	18.3	23	1	GLMS_CLOCO
38	20	18.3	23	1	IRBP_RABIT
39	20	18.3	24	1	ALKB_SALTY
40	20	18.3	24	1	COXC_THUOB
41	20	18.3	24	1	IRBP_SHEEP
42	19.5	17.9	20	1	CAOS_RAT
43	19.5	17.9	25	1	BLP4_BOMOR
44	19	17.4	10	1	AKHX_LOCMI
45	19	17.4	10	1	VEG6_BACSU

ALIGNMENTS

RESULT 1					
KPPI_SELMI					
ID KPPI_SELMI	STANDARD;	PRT;	14 AA.		
AC P25933;					
DT 01-MAY-1992 (Rel. 22, Created)					
DT 01-MAY-1992 (Rel. 22, Last sequence update)					
DT 01-MAY-1992 (Rel. 22, Last annotation update)					
DE PHOSPHORIBULOKINASE, 40 KDA SUBUNIT (EC 2.7.1.19) (PHOSPHOPENTOKINASE)					
DE (FRAGMENT).					
OS Selenastrum minutum.					
OC Eukaryota: Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;					
OC Selenastraceae; Selenastrum.					
RN [1]					
RP SEQUENCE.					
RA Lin M., Turpin D.H.;					
RT "Purification and molecular and immunological characterization of a					
RL unique phosphoribulokinase from the green alga Selenastrum minutum.";					
CC Plant Physiol. 98:82-88(1992).					
CC -1- CATALYTIC ACTIVITY: ATP + D-RIBULOSE 5-PHOSPHATE - ADP +					
CC D-RIBULOSE 1,5-BISPHOSPHATE.					
CC -1- PATHWAY: CALVIN CYCLE.					
CC -1- SUBUNIT: HETERODIMER OF A 40 KDA AND A 41 KDA SUBUNIT.					
CC -1- SIMILARITY: BELONGS TO THE PHOSPHORIBULOKINASE FAMILY.					
DR INTERPRO: IPR001324; -					
DR PROSITE: PS00567; PHOSPHORIBULOKINASE; PARTIAL.					
KW Transferase; Kinase; Calvin cycle; ATP-binding.					
FT NON_TER 14					
SQ SEQUENCE 14 AA; 1379 MW; C4B1D8CD2F891062 CRC64;					
Query Match	26.6%;	Score 29;	DB 1;	Length 14;	
Best Local Similarity	85.7%;	Pred. No. 73;			
Matches 6;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;	
OY 13 MYGLAAD 19					
DB 6 VYGLAAD 12					
RESULT 2					
KPPI_SELMI					
ID KPPI_SELMI	STANDARD;	PRT;	15 AA.		
AC P25934;					
DT 01-MAY-1992 (Rel. 22, Created)					
DT 01-MAY-1992 (Rel. 22, Last sequence update)					
DT 01-MAY-1992 (Rel. 22, Last annotation update)					
DE PHOSPHORIBULOKINASE, 41 KDA SUBUNIT (EC 2.7.1.19) (PHOSPHOPENTOKINASE)					
DE (FRAGMENT).					
OS Selenastrum minutum.					
OC Eukaryota: Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;					
OC Selenastraceae; Selenastrum.					
RN [1]					
RP SEQUENCE.					
RA Lin M., Turpin D.H.;					
RT "Purification and molecular and immunological characterization of a					
RL unique phosphoribulokinase from the green alga Selenastrum minutum.";					

RL Plant Physiol. 98:82-88(1992).
 CC -1- CARLYTIC ACTIVITY: ATP + D-RIBULOSE 5-PHOSPHATE - ADP +
 CC D-RIBULOSE 1,5-BISPHOSPHATE.
 CC -1- PARTWAY: CALVIN CYCLE.
 CC -1- SUBUNIT: HETERODIMER OF A 40 KDA AND A 41 KDA SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHORIBULOKINASE FAMILY.
 DR INTERPRO: IPR001324; -
 DR PROSITE: PS00567; PHOSPHORIBULOKINASE; PARTIAL.
 KW Transferrase; Kinase; Calvin cycle; ATP-binding.
 FT NON_TER 15 15
 SQ SEQUENCE 15 AA; 1455 MW; 07A97BCD2F922C75 CRC64;

Query Match 25.7%; Score 28; DB 1; Length 15;
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 13 MVLGLAD 19
 Db 7 VIGLAD 13

RESULT 3
 RANR_RANRU STANDARD; PRT; 17 AA.
 AC P08952;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE RANATENSIN-R.
 OS Rana rugosa (Frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-84131098; PubMed-6141890;
 RL Nakajima T.;
 RL Unpublished results, cited by:
 RL Erspamer V., Erspamer G.F., Mazzanti G., Endean R.;
 RL Comp. Biochem. Physiol. 77C:99-108(1984).
 CC -1- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
 CC FAMILY.
 DR INTERPRO: IPR000874; -
 DR PFAM: PF02044; Bombesin; 1.
 DR PROSITE: PS00257; BOMBESIN; 1.
 KW Bombesin family; Amidation.
 FT MOD_RES 17 17
 FT SEQUENCE 17 AA; 2053 MW; 3A876B35A581863E CRC64;

Query Match 24.8%; Score 27; DB 1; Length 17;
 Best Local Similarity 55.6%; Pred. No. 1.9e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 RYFNOLSTG 10
 Db 6 RRYNOWATG 14

RESULT 4
 LE01_BIOGL STANDARD; PRT; 25 AA.
 ID LE01_BIOGL
 AC P80742;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HEMOLYMPH 65 KDA LECTIN BG01 (FRAGMENT).
 GN BG01.
 OS Biomphalaria glabrata (Bloodfluke planorb).
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 CC Planorbidae; Biomphalaria.
 RN [1]
 RP SEQUENCE.

RC STRAIN-W-LINE; TISSUE-HEMOLYMPH;
 RX MEDLINE-97385165; PubMed-9238039;
 RA Adema C.M., Hertel L.A., Miller R.D., Loker E.S.;
 RA "A family of fibriogen-related proteins that precipitates parasite-
 RT derived molecules is produced by an invertebrate after infection";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:8691-8696(1997).
 CC -1- FUNCTION: BINDS AND PRECIPITATES ANTIGENS OF THE PARASITE
 CC ECHINOSTOMA PARANSEI.
 CC -1- SUBCELLULAR LOCATION: SECRETED IN THE HEMOLYMPH.
 CC -1- INDUCTION: BY INFECTION.
 CC Lectin.
 KW NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 2602 MW; CC3747240B08810B CRC64;

Query Match 23.9%; Score 26; DB 1; Length 25;
 Best Local Similarity 38.5%; Pred. No. 4.2e+02;
 Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 7 LSTGLDMVGLAD 19
 Db 5 LASGLEMODTTD 17

RESULT 5
 CPAT_PAPSP STANDARD; PRT; 20 AA.
 ID CPAT_PAPSP
 AC P80055;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CYTOCHROME P450 2A7 (EC 1.14.14.1) (CYPI1A7) (P450 F1) (CUMARIN 7-
 DE HYDROXYLASE) (FRAGMENT).
 GN CYP2A7.
 OS Papio sp. (Baboon).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecinae; Papio.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-LIVER;
 RX MEDLINE-92174920; PubMed-1541278;
 RA Dalet-Beluche I., Boulenc X., Fabre G., Maurel P., Bonfils C.;
 RA "Purification of two cytochrome P450 isozymes related to CYP2A and
 RT CYP3A gene families from monkey (baboon, Papio papio) liver
 RT microsomes. Cross reactivity with human forms.";
 RL Eur. J. Biochem. 204:641-648(1992).
 CC -1- FUNCTION: EXHIBITS A HIGH CUMARIN 7-HYDROXYLASE ACTIVITY.
 CC -1- FUNCTION: EXHIBITS A HIGH CUMARIN 7-HYDROXYLASE ACTIVITY.
 CC -1- CARLYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) -> ROH +
 CC OXIDIZED FLAVOPROTEIN + H(2)O.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
 CC -1- INDUCTION: BY PHENOBARBITAL.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR PIR: S21737; S21737.
 DR INTERPRO: IPR001128; -
 DR PROSITE: PS00086; CYTOCHROME_P450; PARTIAL.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT UNSURE 14 14
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2045 MW; 693102A1F0B50C96 CRC64;

Query Match 22.9%; Score 25; DB 1; Length 20;
 Best Local Similarity 50.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 7 LSTGLDMVGLAA 18
 Db 2 LASGLLVALLA 13


```

RESULT 6
UBLI_BOVIN STANDARD; PRT; 25 AA.
AC P23356;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE ISOZYME L1 (EC 3.1.2.15) (UCH-
L1) (UBIQUITIN THIOLESTERASE L1) (NEURON CYTOPLASMIC PROTEIN 9.5)
DE (Pep 9.5) (FRAGMENT).
GN UCHL1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE.
RC TISSUE-BRAIN:
RX MEDLINE-92008646; PubMed-1833240;
RA Giambanco I., Bianchi R., Ceccarelli P., Pula G., Sorci G.,
RA Antonoli S., Bocchini V., Donato R.;
RT "Neuron-specific" protein gene product 9.5 (p9.5) is also
RT expressed in glioma cell lines and its expression depends on cellular
RT growth state."
RT FEBS Lett. 290:131-134(1991).
RL -1- FUNCTION: UBIQUITIN-PROTEIN HYDROLASE IS INVOLVED BOTH IN THE
CC PROCESSING OF UBIQUITIN PRECURSORS AND OF UBIQUINATED PROTEINS.
CC THIS ENZYME IS A THIOL PROTEASE THAT RECOGNIZE AND HYDROLYZE
CC A PEPTIDE BOND AT THE C-TERMINAL GLYCINE OF UBIQUITIN.
CC -1- CATALYTIC ACTIVITY: UBIQUITIN C-TERMINAL THIOLESTER + H(2)O =
CC UBIQUITIN + A THIOL.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: NEURONS AND CELLS OF THE DIFFUSE
CC -1- NEUROENDOCRINE SYSTEM AND THEIR TUMORS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C12; ALSO KNOWN AS
CC PIR: S17561; S17561.
CC DR MEROPS: C12.001.
CC DR INTERPRO: IPR001578;
CC DR PROSITE: PS00140; UCH_1; PARTIAL.
CC FT Ubiquitin conjugation; Hydrolase; Thiol protease; Multigene family.
CC NON_TER 25
CC SEQUENCE 25 AA; 2812 MW; 26B5ADD0A754D55 CRC64;
SQ
Query Match 22.9%; Score 25; DB 1; Length 25;
Best Local Similarity 33.3%; Pred. No. 6.1e+02;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 1 PRYFNQSTGIDMVG 15
DB 10 PEMLNKVLTRLGVAG 24
RESULT 7
ALIS_MANSE STANDARD; PRT; 15 AA.
AC P42539;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ALIATOSTATIN (MAS-AS).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Plekaryota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
RN [1]
RP SEQUENCE.
RC TISSUE-HEAD:
RX MEDLINE-92052112; PubMed-1946359;
RA Kramer S.J., Toschl A., Miller C.A., Kataoka H., Quistad G.B.,
RA Li J.P., Carney R.L., Schooley D.A.;
RT "Identification of an allatostatin from the tobacco hornworm Manduca

```

```

RT sexta."
RL Proc. Natl. Acad. Sci. U.S.A. 88:9458-9462(1991).
CC -1- FUNCTION: STRONGLY INHIBITS JUVENILE HORMONE BIOSYNTHESIS IN VITRO
CC BY THE CORPORA ALATA FROM FIFTH-STADIUM LARVAE AND ADULT FEMALES.
CC -1- SIMILARITY: BELONGS TO THE ALIATOSTATIN FAMILY.
CC KW Neuropeptide.
FT MOD_RES 1
FT MOD_RES 1 PYRROLIDONE CARBOXYLIC ACID.
FT SEQUENCE 15 AA; 1908 MW; 1605B77CDBEC838E CRC64;
SQ
Query Match 22.0%; Score 24; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 YFNQIS 8
DB 8 YFNPIIS 13
RESULT 8
OXFE_SCYCA STANDARD; PRT; 9 AA.
ID OXFE_SCYCA
AC P42997;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PHASVATOICIN.
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.
RN [1]
RP SEQUENCE.
RC TISSUE-PITUITARY;
RX MEDLINE-95062247; PubMed-7972045;
RA Chauvet J., Rouille Y., Chauveau C., Chauvet M.-T., Acher R.;
RT "Special evolution of neurohypophysial hormones in cartilaginous
RT fishes: aspartocin and phasvatoicin, two oxytocin-like peptides
RT isolated from the spotted dogfish (Scyllorhinus caniculus).";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11266-11270(1994).
CC -1- FUNCTION: DISPLAYS OXYTOCIC ACTIVITY ON RAT UTERUS.
CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
CC DR INTERPRO: IPR000981;
CC DR PFMW: PF00220; hormone; 1.
CC DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
CC FT Hormone; Amidation.
CC FT DISULFID 1
CC FT MOD_RES 9
CC SEQUENCE 9 AA; 1016 MW; 17EDD76EB44449DB CRC64;
SQ
Query Match 21.1%; Score 23; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 8.9e+04;
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 YFNQISG 10
DB 2 YFNCPVG 9
RESULT 9
PPBH_PSEAE STANDARD; PRT; 19 AA.
ID PPBH_PSEAE
AC P35483;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALKALINE PHOSPHATASE H (EC 3.1.3.1) (H-AP) (FRAGMENT).
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
RN [1]
RP SEQUENCE.

```

```

RC STRAIN-H103;
RX MEDLINE-93202452; PubMed-8454193;
RA Tan A.S.P., Morobac E.A.;
RT "Isolation and characterization of two immunochemically distinct
  alkaline phosphatases from Pseudomonas aeruginosa."
RL FEMS Microbiol. Lett. 106:281-286(1993).
CC -1- FUNCTION: HAS ONLY A PHOSPHOMONESTERASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
CC ALCOHOL + ORTHOPHOSPHATE (AT A HIGH PH OPTIMUM).
CC -1- COFACTOR: METALLOENZYME CONTAINING TWO ZINC ATOMS AND A MAGNESIUM
CC ION.
CC -1- SUBCELLULAR LOCATION: SECRETED AND PERIPLASMIC.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED CONSTITUTIVELY.
CC -1- SIMILARITY: BELONGS TO THE ALKALINE PHOSPHATASE FAMILY.
DR INTERPRO: IPR001952;
DR PROSITE: PS00123; ALKALINE_PHOSPHATASE; PARTIAL.
KW Hydrolase; zinc; Magnesium; Periplasmic.
FT NON_TER 19
SQ SEQUENCE 19 AA; 2131 MW; C51B09D7DB22E799 CRC64;

Query Match
Best Local Similarity 21.1%; Score 23; DB 1; Length 19;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PRYFNOLSTG 10
DB 5 PSLFNROAG 14

RESULT 10
ID OMPL_ACTAC STANDARD; PRT; 20 AA.
AC P20242;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE 39 kDa MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
OS Actinobacillus actinomycetemcomitans (Haemophilus
  actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
  Actinobacillus.
RN [1]
RP SEQUENCE.
RC STRAIN-Y4;
RX MEDLINE-91169244; PubMed-2004699;
RA Kokeguchi S., Kato K., Nishimura F., Kurihara H., Murayama Y.;
RT "Isolation and partial characterization of a 39 kDa major outer
  membrane protein of Actinobacillus actinomycetemcomitans Y4."
RL FEMS Microbiol. Lett. 61:85-89(1991).
DR PIR: A54538; A54538.
KW Outer membrane; Transmembrane; Porin.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2316 MW; A837A8C4764F527E CRC64;

Query Match
Best Local Similarity 21.1%; Score 23; DB 1; Length 20;
Matches 3; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 RYFNOLSTGLDM 13
DB 2 KVTNOSTKVEL 13

RESULT 11
ID KAD_BACLI STANDARD; PRT; 24 AA.
AC P35140;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSNOSHOPHYLASE) (FRAGMENT).

```

```

GN ADK.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93062802; PubMed-1435726;
RA Tschander S., Driessen A.J.M., Freudl R.;
RT "Cloning and molecular characterization of the sec genes from
  Bacillus licheniformis and Staphylococcus carnosus: comparative
  analysis of nine members of the SecY family."
RL Mol. Gen. Genet. 235:147-152(1992).
CC -1- FUNCTION: THIS SMALL UNICITOUS ENZYME IS ESSENTIAL FOR
CC MAINTENANCE AND CELL GROWTH.
CC -1- CATALYTIC ACTIVITY: ATP + AMP -> ADP + ADP.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE ADENYLATE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X70087; -; NOT_ANNOTATED_CDS.
CC PIR: S34405; S34405.
CC HSSP: P27142; 1Z10.
CC INTERPRO: IPR000650;
CC PIR: P00406; adenylatekinase; 1.
CC PROSITE: PS00113; ADENYLATE_KINASE; PARTIAL.
CC TRANSFERASE; Kinase; ATP-binding.
FT NP_BIND 7
FT NON_TER 24
SQ SEQUENCE 24 AA; 2563 MW; A65B1A5F1B018F21 CRC64;

Query Match
Best Local Similarity 21.1%; Score 23; DB 1; Length 24;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 LDMVGL 16
DB 3 LDMGL 8

RESULT 12
ID PGO_XENIA STANDARD; PRT; 24 AA.
AC P39080;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE ANTIMICROBIAL PEPTIDE PGO.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE.
RC TISSUE-STOMACH;
RX MEDLINE-92011794; PubMed-1717472;
RA Moore K.S., Bevins C.L., Brasseur M.M., Tomassini N., Turner K.,
  Eck H., Zasloff M.;
RT "Antimicrobial peptides in the stomach of Xenopus laevis."
RL J. Biol. Chem. 266:19851-19857(1991).
CC -1- FUNCTION: ANTIMICROBIAL PEPTIDE.
CC -1- TISSUE SPECIFICITY: IS SYNTHESIZED IN THE STOMACH AND STORED
CC IN A NOVEL GRANULAR MULTINUCLEATED CELL IN THE GASTRIC MUCOSA.
CC IT IS STORED AS ACTIVE, PROCESSED PEPTIDES IN LARGE GRANULES
CC WITHIN THE GRANULAR GLAND SECRETIONS OF THE SKIN.

```

CC -1- SIMILARITY: BELONGS TO THE MAGAININ FAMILY OF ANTIMICROBIAL
 CC PEPTIDES.
 DR PIR: A41037; A41037.
 KW Antibiotic; Amphibian skin.
 SQ SEQUENCE 24 AA; 2457 MW; 7E6A87CB7CE22B9C CRC64;

Query Match 21.1%; Score 23; DB 1; Length 24;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 YFNOLSTG 10
 | : | | |
 Db 9 YLRKLG 16

RESULT 13
 CR21_LITSP STANDARD; PRT; 25 AA.
 AC P56233;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE CAERIN 2.1.
 OS Litoria splendida.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 CC Litoria.
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE-PAROTOID GLAND;
 RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
 RT "Peptides from Australian frogs. Structures of the caerins and
 ceradgin 1 from Litoria splendida."
 RL J. Chem. Soc. Perkin Trans. 1:3173-3178(1992).
 CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE THAT ADOPTS AN ALPHA HELICAL
 CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
 CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
 CC GLANDS.
 KW Antibiotic; Amphibian skin.
 KM Antibiologic; Amphibian skin.
 SQ SEQUENCE 25 AA; 2394 MW; DDCA9BC6B49186B8 CRC64;

Query Match 21.1%; Score 23; DB 1; Length 25;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 8 STGDMVGLAD 19
 | : | | |
 Db 5 SIGRAGGLAD 16

RESULT 14
 CR22_LITGI STANDARD; PRT; 25 AA.
 AC P56234;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE CAERIN 2.2 [CONTAINS: CAERIN 2.2.1].
 OS Litoria gilleni, and Litoria caerulea.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 CC Litoria.
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY OF CAERIN 2.2.
 RC SPECIES-L. CAERULEA; TISSUE-PAROTOID GLAND;
 RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
 RT "Peptides from Australian frogs. The structures of the caerins from
 Litoria caerulea."
 RL J. Chem. Res. 138:910-936(1993).

RN [2]
 RP SEQUENCE, AND MASS SPECTROMETRY OF CAERIN 2.2.1.
 RC SPECIES-L. GILLIENI; TISSUE-PAROTOID GLAND;
 RA Maugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;
 RT "Peptides from Australian frogs. The structures of the caerins and
 ceradgins from Litoria gilleni."
 RL J. Chem. Res. 139:937-961(1993).
 CC -1- FUNCTION: ANTIMICROBIAL PEPTIDE THAT ADOPTS AN ALPHA HELICAL
 CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
 CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
 CC GLANDS.
 CC -1- MASS SPECTROMETRY: MM=2464; METHOD-FAB; RANGE=1-25.
 CC -1- MASS SPECTROMETRY: MM=1695; METHOD-FAB; RANGE=9-25.
 KW Antibiotic; Amphibian skin.
 FT PEPTIDE 1 25 CAERIN 2.2.
 FT PEPTIDE 9 25 CAERIN 2.2.1.
 SQ SEQUENCE 25 AA; 2466 MW; DDCA9BC5D49186B8 CRC64;

Query Match 21.1%; Score 23; DB 1; Length 25;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 8 STGDMVGLAD 19
 | : | | |
 Db 5 SIGRAGGLAD 16

RESULT 15
 CR23_LITCE STANDARD; PRT; 25 AA.
 ID CR23_LITCE
 AC P56235; P82117;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SPENDIPHERIN (CAERIN 2.3).
 OS Litoria caerulea, and
 DE Litoria splendida.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 CC Litoria.
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC SPECIES-L. CAERULEA; TISSUE-PAROTOID GLAND;
 RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
 RT "Peptides from Australian frogs. The structures of the caerins from
 Litoria caerulea."
 RL J. Chem. Res. 138:910-936(1993).
 RN [2]
 RP SEQUENCE.
 RC SPECIES-L. SPLENDIDA; TISSUE-SKIN SECRETION;
 RX MEDLINE-99447035; PubMed-10519546;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C., Smith B.P.;
 RT "Aquatic sex pheromone from a male tree frog."
 RL Nature 401:444-445(1999).
 RN [3]
 RP SEQUENCE.
 RC SPECIES-L. SPLENDIDA; TISSUE-SKIN SECRETION;
 RX MEDLINE-20069371; PubMed-10601876;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C., Smith B.P.;
 RT "Differences in the skin peptides of the male and female Australian
 tree frog Litoria splendida. The discovery of the aquatic male sex
 pheromone splendipherin, together with phed caerulein and the
 antidiabetic peptide caerin 1.20."
 RL Eur. J. Biochem. 267:269-275(2000).
 CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE THAT ADOPTS AN ALPHA HELICAL
 CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
 CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
 CC -1- FUNCTION: ACTS AS A MALE SEX PHEROMONE THAT ATTRACTS FEMALE.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
 CC GLANDS.
 CC -1- MASS SPECTROMETRY: MM=2364; METHOD-FAB.

KW Antibiotic; Pheromone; Amphibian skin.
SQ SEQUENCE 25 AA; 2366 MW; DDD82C36B49186B8 CRC64;

Query Match 21.1%; Score 23; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 8 STGLDMVGLAAD 19
1 1 : 1 1 1
Db 5 SIGKALGGLAD 16

RESULT 16
ID CR24.LITCE STANDARD; PRT; 25 AA.
AC P56236;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CAERIN 2.4.
OS Litoria caerulea.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-PAROTOID GLAND;
RA Stone D.J.M., Maugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins from
RT Litoria caerulea."
J. Chem. Res. 138:910-936(1993).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL
CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC GLANDS.
CC -1- MASS SPECTROMETRY: MW=2450; METHOD-FAB.
CC Antibiotic; Amphibian skin.
SQ SEQUENCE 25 AA; 2452 MW; DDD82C35D49596B8 CRC64;

Query Match 21.1%; Score 23; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 8 STGLDMVGLAAD 19
1 1 : 1 1 1
Db 5 SIGKALGGLAD 16

RESULT 17
ID CR25.LITGI STANDARD; PRT; 25 AA.
AC P56237;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CAERIN 2.5.
OS Litoria gilleni.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-PAROTOID GLAND;
RA Maugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins and
RT caeridins from Litoria gilleni."
J. Chem. Res. 139:937-961(1993).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL
CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.

CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC GLANDS.
CC -1- MASS SPECTROMETRY: MW=2448; METHOD-FAB.
KW Antibiotic; Amphibian skin.
SQ SEQUENCE 25 AA; 2450 MW; DDCA9BC5D48F8758 CRC64;

Query Match 21.1%; Score 23; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 8 STGLDMVGLAAD 19
1 1 : 1 1 1
Db 5 SIGKALGGLAD 16

RESULT 18
ID CBPR_PROAT STANDARD; PRT; 15 AA.
AC P19628;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE CARBOXYPEPTIDASE B (EC 3.4.17.2) (FRAGMENT).
OS Proteolus aethiopicus (Marbled lungfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Dipnoi; Lepidosireniformes; Protopterygidae; Protopterus.
RN [1]
RP SEQUENCE.
RC TISSUE-PANCREAS;
RX MEDLINE-73025047; PubMed-5079891;
RA Reek G.R., Neurath H.;
RT "Isolation and characterization of pancreatic procarboxypeptidase B
RT and carboxypeptidase B of the African lungfish."
J. Biochemistry 11:3947-3955(1972).
CC -1- CATALYTIC ACTIVITY: PEPTIDYL-L-LYSINE/ARGININE + H(2)O - PEPTIDE +
CC L-LYSINE/ARGININE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M14; ALSO KNOWN AS THE
CC ZINC CARBOXYPEPTIDASE FAMILY.
CC PIR: A26212; A26212.
DR MEROPS: M14.003; -;
DR INTERPRO: IPR000834; -;
DR PROSITE: PS00132; CARBOXYPEPT_ZN_1; PARTIAL.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; PARTIAL.
KW Hydrolase; Carboxypeptidase; Zinc; Zymogen.
FT PROPEP 1
FT NON_TER 15
FT SEQUENCE 15 AA; 1749 MW; 124C910D937BED65 CRC64;

Query Match 20.2%; Score 22; DB 1; Length 15;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PRYFN 5
1 1 : 1 1
Db 4 PRSFN 8

RESULT 19
ID OXLA.OPHHA STANDARD; PRT; 19 AA.
AC P81383;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE L-AMINO ACID OXIDASE (EC 1.4.3.2) (LAO) (LAAO) (FRAGMENT).
OS Ophiophagus hannah (King cobra) (Naja hannah).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Elapinae; Ophiophagus.
RN [1]
RP SEQUENCE.

RC TISSUE-VENOM; PubMed-8080286;
 RX MEDLINE-94361525; PubMed-8080286;
 RA Ponnudurai G., Chung M.C.M., Tan N.-H.;
 RT "Purification and properties of the L-amino acid oxidase from Malaysian
 RT pit viper (Calloselasma rhodostoma) venom.";
 RL Arch. Biochem. Biophys. 313:373-378(1994).
 RN [2]
 RP SEQUENCE OF 1-15.
 RC TISSUE-VENOM;
 RX MEDLINE-97449790; PubMed-9304806;
 RA Ahn M.Y., Lee B.M., Kim Y.S.;
 RT "Characterization and cytotoxicity of L-amino acid oxidase from the
 RT venom of King cobra (Ophiophagus hannah).";
 RL Int. J. Biochem. Cell Biol. 29:911-919(1997).
 CC -1- FUNCTION: HAS CYTOTOXIC ACTIVITY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: A L-AMINO ACID + H(2)O + O(2) -> A 2-OXO ACID +
 CC NH(3) + H(2)O(2).
 CC -1- COFACTOR: FAD.
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).
 CC -1- PTM: GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOXAMINE OXIDASE FAMILY.
 CC STRONG, TO MOUSE FIG-1.
 KM Oxidoreductase: Flavoprotein; FAD; Glycoprotein; Venom.
 FT CONFLICT 1 1 H -> S (IN REF. 2).
 FT NON_TER 19 19
 SQ SEQUENCE 19 AA; 2298 MW; DD911A5B414F1427 CRC64;

Query Match 20.28; Score 22; DB 1; Length 19;
 Best Local Similarity 50.08; Pred. No. 1.4e+03;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PYRENOLS 8
 DB 12 PEYENHLLA 19

RESULT 20
 FIBR_PACLE STANDARD; PRT; 20 AA.
 AC P81070;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE FIBRINOGEN (VHDL) (FRAGMENT).
 OS Pacificastacus leniusculus (Signal crayfish).
 OS Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
 OC Astacoidae; Astacidae; Pacifastacus.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-96074573; PubMed-7488215;
 RA Hall M., van Heusden M.C., Soederhaell K.;
 RT "Identification of the major lipoproteins in crayfish hemolymph as
 RT proteins involved in immune recognition and clotting.";
 RL Biochem. Biophys. Res. Commun. 216:939-946(1995).
 CC -1- FUNCTION: INVOLVED IN LIPID TRANSPORT. PLAYS A ROLE IN HEMOLYMPH
 CC CLOTTING. MAY BE INVOLVED IN WOUND HEALING IN THE CUTICLE.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- TISSUE SPECIFICITY: SECRETED INTO THE HEMOLYMPH.
 CC -1- SIMILARITY: TO VITELLOGENINS.
 KM Coagulation; Glycoprotein; Lipid-binding; Hemolymph.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2311 MW; 1C04BC596B9FC47 CRC64;

Query Match 20.28; Score 22; DB 1; Length 20;
 Best Local Similarity 30.08; Pred. No. 1.5e+03;
 Matches 3; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 2 RYFNOLSTGL 11
 DB 10 RYSGRVAAGI 19

RESULT 21
 SOD2_PICAB STANDARD; PRT; 21 AA.
 AC P29428;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SUPEROXIDE DISMUTASE [CU-ZN] II (EC 1.15.1.1) (SOD II) (FRAGMENT).
 OS Picea abies (Norway spruce) (Picea excelsa).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Coniferopsida; Coniferales; Pinaceae; Picea.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-NEEDLE;
 RA Koenigler W., Rensenberg H., Polle A.;
 RT "Purification of two superoxide dismutase isozymes and their
 RT subcellular localization in needles and roots of Norway spruce (Picea
 RT abies L.) trees.";
 RL Plant Physiol. 100:334-340(1992).
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
 CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) -> O(2) + H(2)O(2).
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- TISSUE SPECIFICITY: DOMINANT ISOZYME IN ROOTS.
 CC -1- MISCELLANEOUS: EUKARYOTIC CELLS CONTAIN A MITOCHONDRIAL
 CC MN-CONTAINING ENZYME & A CYTOPLASMIC CU-ZN-CONTAINING ENZYME.
 CC PLANTS ALSO HAVE A CHLOROPLAST CU-ZN ENZYME.
 CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 CC HSP; P15107; 1XSO.
 DR INTERPRO: IPR001424;
 DR PFAM: PF00080; soccu; 1.
 DR PROSITE: PS00087; SOD_CU_ZN_1; PARTIAL.
 DR PROSITE: PS00332; SOD_CU_ZN_2; PARTIAL.
 KM Oxidoreductase; Copper; Zinc; Multigene family.
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2100 MW; 6CF0E108CAF03D65 CRC64;

Query Match 20.28; Score 22; DB 1; Length 21;
 Best Local Similarity 50.08; Pred. No. 1.6e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 TGLDMVGL 16
 DB 10 TGADEVKV 17

RESULT 22
 ALL4_OLEEU STANDARD; PRT; 24 AA.
 AC P80741;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE MAJOR POLLEN ALLERGEN OLE E 4 (OLE E IV) (FRAGMENTS).
 OS Olea europaea (Common olive).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Lamiales; Oleaceae; Olea.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-POLLEN;
 RX MEDLINE-98160390; PubMed-9500754;
 RA Boluda L., Alonso C., Fernandez-Caldas E.;
 RT "Purification, characterization, and partial sequencing of two new
 RT allergens of Olea europaea".
 RL J. Allergy Clin. Immunol. 101:210-216(1998).
 CC -1- DISEASE: MAJOR ALLERGEN FROM OLIVE POLLEN. IMPORTANT IN
 CC MEDITERRANEAN COUNTRIES (BY SIMILARITY).
 KM Allergen.

FT NON_TER 1 1
 RT NON_CONS 10 11
 FT NON_TER 24 24
 SQ SEQUENCE 24 AA; 2711 MW; ADAC5DA9F577D6D CRC64;

Query Match
 Best Local Similarity 33.3%; Score 22; DB 1; Length 24;
 Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 8 STGLDMVGL 16
 DB 4 MTGELVSI 12

RESULT 23
 GAE6_RANRU STANDARD; PRT; 24 AA.

AC P80400;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE GAEGRIN-6.
 OS Rana rugosa (Frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=SKIN;
 RX MEDLINE=95091844; PubMed=7999137;
 RA Park J.M., Jung J.-E., Lee B.J.;
 RT "Antimicrobial peptides from the skin of a Korean frog, Rana rugosa."
 RL Biochem. Biophys. Res. Commun. 205:948-954(1994).
 CC -1- FUNCTION: HAS A NON-HEMOLYTIC ACTIVITY. HAS A BROAD SPECTRUM OF ACTIVITY AGAINST BOTH GRAM POSITIVE AND GRAM NEGATIVE BACTERIA, FUNGI AND PROTOZOA.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: SKIN.
 CC -1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN FAMILY.
 KW Amphibian skin; Antiploetic.
 FT DISULFID 18 24
 SQ SEQUENCE 24 AA; 2610 MW; 09918123FF90CCD CRC64;

Query Match
 Best Local Similarity 50.0%; Score 22; DB 1; Length 24;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 13 MVGLADW 20
 DB 5 LAGLANF 12

RESULT 24
 NODD_RHILLO STANDARD; PRT; 18 AA.
 AC 052838;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE NODULATION PROTEIN D (FRAGMENT).
 OS Rhizobium loti (Mesorhizobium loti).
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Phyllobacteriaceae; Mesorhizobium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NZP 2213;
 RX MEDLINE=97002748; PubMed=8550088;
 RA Scott D.B., Young C.A., Collins-Emerison J.M., Terzaghi E.A.,

RA Rockman E.S., Lewis P.E., Pankhurst C.E.;
 RT "Novel and complex chromosomal arrangement of Rhizobium loti nodulation genes."
 RL Mol. Plant Microbe Interact. 9:187-197(1996).
 CC -1- FUNCTION: NODD REGULATES THE EXPRESSION OF THE NOD ABCF GENES WHICH ENCODE OTHER NODULATION PROTEINS. NODD IS ALSO A NEGATIVE REGULATOR OF ITS OWN EXPRESSION. BINDS FLAVENOID AS INDUCERS.
 CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L06241; AAB47350.1; -
 DR INTERPRO; IPR000847; -
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; PARTIAL.
 KW Nodulation; Transcription regulation; DNA-binding; Activator; Repressor.
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 2170 MW; 89BA8A62B591136 CRC64;

Query Match
 Best Local Similarity 41.7%; Score 21.5; DB 1; Length 18;
 Matches 5; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

OY 10 GLDM-VGLADW 20
 DB 5 GDLNLLVPPDM 16

RESULT 25
 CN2_LITGI STANDARD; PRT; 15 AA.
 AC P56247;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE CAERIDIN 2.
 OS Litorea gilleni.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae; Liorea.
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=PAROTOID GLAND;
 RA Waugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;
 RT "Peptides from Australian frogs. The structures of the caerins and caeridins from Litorea gilleni."
 RL J. Chem. Res. 139:937-961(1993).
 CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTI-BIOTIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR POSTAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=1408; METHOD=FAE.
 KW Amphibian skin; Amidation.
 FT MOD_RES 15 15
 SQ SEQUENCE 15 AA; 1410 MW; 06F1BBF72550CBF CRC64;

Query Match
 Best Local Similarity 80.0%; Score 21; DB 1; Length 15;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 11 LDMVG 15
 DB 3 LDVVG 7

Tue Feb 6 08:48:22 2001

Search completed: February 5, 2001, 10:56:02
Job time: 502 sec

us-08-981-824-5.rsp

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:55:26 ; Search time 93.77 Seconds
(without alignments)
24.999 Million cell updates/sec

Title: US-08-981-824-5
Sequence: 109
1 PRYFNQLSTGLDMVGLADW 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP rhiz:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	31.2	11	6	Q9TRX0
2	30	27.5	20	6	Q9TRB9
3	27	24.8	24	11	Q89021
4	27	24.8	25	9	Q92WY2
5	27	24.8	25	10	Q48860
6	26	23.9	12	6	Q9XT42
7	26	23.9	22	11	Q62538
8	26	23.9	22	2	Q85543
9	25	22.9	15	7	Q9TNQ1
10	25	22.9	15	8	Q99386
11	25	22.9	16	2	Q9RAF2
12	25	22.9	16	2	Q9RSE9
13	25	22.9	19	11	Q9QVJ9
14	25	22.9	22	2	Q47529
15	25	22.9	22	10	Q48861
16	25	22.9	23	8	Q9TJ33
17	24	22.0	16	11	Q9QV05
18	24	22.0	23	12	Q69392
19	23	21.1	13	10	Q9S941

20	23	21.1	15	12	Q9PXC5	Q9PXC5 tobacco etc
21	23	21.1	17	4	Q9UCP7	Q9UCP7 homo sapien
22	23	21.1	17	6	Q9TRD9	Q9TRD9 bos taurus
23	23	21.1	17	10	Q24445	Q24445 ceratodon p
24	23	21.1	18	10	Q9S888	Q9S888 narcissus p
25	23	21.1	19	8	Q9TJ32	Q9TJ32 nicotiana t
26	23	21.1	20	10	Q9S903	Q9S903 vigna sinen
27	23	21.1	22	2	P78167	P78167 escherichia
28	23	21.1	22	4	Q16415	Q16415 homo sapien
29	23	21.1	23	13	Q9UWK1	Q9UWK1 methanosarc
30	23	21.1	24	10	Q9S890	Q9S890 sambucus n1
31	23	21.1	24	12	Q73766	Q73766 human immun
32	23	21.1	25	5	Q26087	Q26087 polycells n
33	22.5	20.6	23	13	P82400	P82400 litorea ran
34	22	20.2	9	13	P82075	P82075 litorea rub
35	22	20.2	9	13	P82093	P82093 litorea rub
36	22	20.2	12	2	Q9R7E1	Q9R7E1 staphylococ
37	22	20.2	12	8	P92454	P92454 cycas revol
38	22	20.2	13	2	P94878	P94878 lactococcus
39	22	20.2	13	10	Q43174	Q43174 solanum tub
40	22	20.2	15	1	Q9UWM1	Q9UWM1 methanospir
41	22	20.2	16	4	Q9UD41	Q9UD41 homo sapien
42	22	20.2	17	11	P97758	P97758 mus musculu
43	22	20.2	18	11	Q9JIM8	Q9JIM8 rattus norv
44	22	20.2	19	2	Q9R5H4	Q9R5H4 rhodospirill
45	22	20.2	20	2	Q9R4Y5	Q9R4Y5 heliobacte

ALIGNMENTS

RESULT 1
Q9TRX0 PRELIMINARY; PRT; 11 AA.
AC Q9TRX0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE LANOSTEROL 14 ALPHA-DEMETHYLASE, CYTOCHROME P-45014DM.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
CX NCBI TaxID=9623;
RN [1]
RP MEDLINE=91316123; PubMed=1859829;
RX Sono H., Sonoda Y., Sato Y.;
RA "Purification and characterization of cytochrome P-45014DM (lanosterol
RT 14 alpha-demethylase) from pig liver microsomes.";
RL Biochim. Biophys. Acta 1078:388-394(1991).
SQ SEQUENCE 11 AA; 1084 MW; 8A7A5C8C2AA7861 CRC64;

Query Match 31.2%; Score 34; DB 6; Length 11;
Best Local Similarity 60.0%; Pred. NO. 30;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 LSTGLDMVGL 16
Db 2 LITGLDLGI 11
RESULT 2
Q9TRB9 PRELIMINARY; PRT; 20 AA.
AC Q9TRB9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ENTEROTOXIN-BINDING GLYCOPROTEIN PP2OK (FRAGMENT).
OS Bos taurus (Bovine).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae; Bovidae;

```

OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RP [1]
RN SEQUENCE.
RX MEDLINE=94259890; PubMed=8201051;
RA Shida K., Takamizawa K., Nagaoka M., Kushihiro A., Osawa T., Tsuji T.,
RT "Enterotoxin-binding glycoproteins in a protease-peptide fraction of
RT heated bovine milk."
RL J. Dairy Sci. 77:930-939(1994).
DR HSP: P02754; IBSY.
SQ SEQUENCE 20 AA; 2266 MW; C8236200247ABA66 CRC64;

Query Match
Best Local Similarity 40.0%; Score 30; DB 6; Length 20;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 6 Q1STGLDMVGLADW 20
DB 5 Q1MKGLDIOKRVAGTW 19

RESULT 3
089021 PRELIMINARY; PRT; 24 AA.
AC 089021;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE MUSCULAR CHLORIDE CHANNEL 1 (FRAGMENT).
GN CLC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-SKELETAL MUSCLE;
RA Vullhorst D.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ011106; CAA09489.1;
FT NON_TER 1
FT SEQUENCE 24 AA; 2502 MW; B1AB84650DC0BF64 CRC64;

Query Match
Best Local Similarity 38.1%; Score 27; DB 11; Length 24;
Matches 8; Conservative 1; Mismatches 6; Indels 6; Gaps 1;

OY 6 Q1STG-----IDMVGGLADW 20
DB 1 Q1KSGPLAGISLGMVCGRAW 21

RESULT 4
092WY2 PRELIMINARY; PRT; 25 AA.
AC 092WY2;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE ILEX, STX2A, STX2B, GENES AND 25 ORF'S.
OS Bacteriophage 933W.
OC Viruses; dsDNA viruses, no RNA stage; unclassified dsDNA phages.
OX NCBI_TaxID=10730;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97076910; PubMed=8975608;
RA Datz M., Janetzki-Mittmann C., Franke S., Gunzer F., Schmidt H.,
RA Karch H.;
RT "Analysis of the enterohemorrhagic Escherichia coli O157 DNA region
RT containing lambdoid phage gene p and Shiga-like toxin structural
RT genes."

```

```

RL Appl. Environ. Microbiol. 62:791-797(1996).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=97257502; PubMed=9103976;
RX Schmidt H., Scheef J., Janetzki-Mittmann C., Datz M., Karch H.;
RT "An ilex cDNA gene is located close to the Shiga toxin II operon in
RT enterohemorrhagic Escherichia coli O157 and non-O157 strains."
RL FEMS Microbiol. Lett. 149:39-44(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Karch H., Schmidt H., Janetzki-Mittmann C., Scheef J., Kroeger M.;
RT "Shiga toxins, even when different are encoded in identical positions
RT in related temperate bacteriophages."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y10775; CAB39286.1;
SQ SEQUENCE 25 AA; 2779 MW; F2504C217C2A905C CRC64;

Query Match
Best Local Similarity 33.3%; Score 27; DB 9; Length 25;
Matches 6; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 3 YFNOLSTGLDMVGLADW 20
DB 4 YTELITGLVINGLEFIW 21

RESULT 5
048860 PRELIMINARY; PRT; 25 AA.
AC 048860;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HOMEOBOX PROTEIN (FRAGMENT).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IRB21;
RA Liu G., Yang J., Zhai W., He P., Li X., Lu J., Li S., Zhu L.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003599; AAB94484.1;
DR INTERPRO; IPR001356;
DR PROSITE; PS50071; HOMEOBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein.
FT NON_TER 1
FT SEQUENCE 25 AA; 2861 MW; CC14641CF1E5BFCE CRC64;

Query Match
Best Local Similarity 83.3%; Score 27; DB 10; Length 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 8 STGLDM 13
DB 10 STGLDL 15

RESULT 6
09XT42 PRELIMINARY; PRT; 12 AA.
AC 09XT42;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE ALPHA 5 CHAIN COLLAGEN TYPE IV (FRAGMENT).
GN COL4A5.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99293036; PubMed=10362815;
RA Zheng K., Harvey S., Sado Y., Naito I., Ninomiya Y., Jacobs R.,
  Thorner P.S.;
RT "Absence of the alpha6(IV) chain of collagen type IV in Alport
  syndrome is related to a failure at the protein assembly level and
  does not result in diffuse glomerulonephritis."
RL Am. J. Pathol. 154:1883-1891(1999).
DR EMBL: AF128530; AAD45500.1; -.
KW Collagen.
FT NON_TER
SQ SEQUENCE 12 AA; 1289 MW; 2FBC5925EE3A2C8 CRC64;

Query Match      23.9%; Score 26; DB 6; Length 12;
Best Local Similarity 71.4%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 14 VGLADW 20
DB 6 VSLAGW 12

RESULT 7
062538 PRELIMINARY; PRT; 21 AA.
AC 062538;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE GRANTYME B(G,H) (EC 3.4.21.79) (CYTOTOXIC CELL PROTEASE 1) (CCP1)
  (CTLA-1) (FRAGMENTIN 2) (FRAGMENT).
GN GZM OR CTLA-1.
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRET/ET;
RX MEDLINE=94319082; PubMed=8043949;
RA Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maezaki Y.,
  Nadeau J.H.;
RT "Genetic mapping of 40 cDNA clones on the mouse genome by PCR."
  Mamm. Genome 5:349-355(1994).
RL
CC -1- FUNCTION: THIS ENZYME IS NECESSARY FOR TARGET CELL LYSIS IN CELL-
  MEDIATED IMMUNE RESPONSES. IT CLEAVES AFTER ASP. SEEMS TO BE
  LINKED TO AN ACTIVATION CASCADE OF CASPASES (ASPARTATE-SPECIFIC
  CYSTEINE PROTEASES) RESPONSIBLE FOR APOPTOSIS EXECUTION. CLEAVES
  CASPASE-3, -7, -9 AND 10 TO GIVE RISE TO ACTIVE ENZYMES MEDIATING
  APOPTOSIS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ASP-1-XAA >> ASN-1-XAA
  > MET-1-XAA, SER-1-XAA.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES OF CYTOLYTIC T-
  LYMPHOCYTES AND NATURAL KILLER CELLS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
  TRYPSIN FAMILY. STRONGEST TO OTHER GRANZYMES AND TO MAST CELL
  PROTEASES.
DR EMBL: U05708; AAB60471.1; -.
DR HSSP: P0187; 2CPL.
DR MEROPS: S01.136; -.
KM MGD: MGI:109267; Gmb.
KW Hydrolyase; Serine protease; Zymogen; T-cell; Cytolysis; Apoptosis.
FT NON_TER
SQ SEQUENCE 21 AA; 2430 MW; 9D4C8253DFB1AE6 CRC64;

Query Match      23.9%; Score 26; DB 11; Length 21;
Best Local Similarity 35.7%; Pred. No. 1.3e+03;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 1 PRYFNOLSTGLDNY 14
DB 1 PRAFTKVSFLSWI 14

RESULT 8
085543 PRELIMINARY; PRT; 22 AA.
AC 085543;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
GN OMP-1.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REF A12B;
RA Hsieh Y.-H., Bobo L.D.;
RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
  trachomatis in trachoma endemic villages, Korea, Tanzania."
  Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF070327; AAC25298.1; -.
DR
FT NON_TER
FT NON_TER
SQ SEQUENCE 22 AA; 2237 MW; 3BD2052033F8097D CRC64;

Query Match      23.9%; Score 26; DB 2; Length 22;
Best Local Similarity 54.5%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 9 TGDVNGSLAD 19
DB 4 TSDVNGLEKD 14

RESULT 9
09TN01 PRELIMINARY; PRT; 15 AA.
AC 09TN01;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE BETA 2M-CLASS I-BINDING PEPTIDE-MAJOR HISTOCOMPATIBILITY COMPLEX
  H-2KB-SPECIFIC MOLECULE POORLY ASSOCIATED WITH BETA 2-MICROGLOBULIN.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE.
RX MEDLINE=94240094; PubMed=8183884;
RA Joyce S., Kuzushima K., Kepecs G., Angelelli R.H., Nathenson S.G.;
RT "Characterization of an incompletely assembled major
  histocompatibility class I molecule (H-2Kb) associated with unusually
  long peptides: implications for antigen processing and presentation."
  Proc. Natl. Acad. Sci. U.S.A. 91:4145-4149(1994).
RL
KM MHC.
SQ SEQUENCE 15 AA; 1754 MW; 2CEACDA0F1180DC7 CRC64;

```

```

OY 1 PRYFNOL 7
DB 1 PRYFNOL 7

Query Match      22.9%; Score 25; DB 7; Length 15;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

RESULT 10
099386 PRELIMINARY; PRT; 15 AA.
AC 099386:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CYTOCHROME OXIDASE SUBUNIT 2 (FRAGMENT).
GN COIT.
OS Sus scrofa (Pig).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PERIPHERAL BLOOD;
RX MEDLINE=98403943; PubMed=9734874;
RA Tartaglia M., Saulle E.;
RT "Rapid communication: nucleotide sequence of porcine and ovine
RL TRN(Lys) and ATPase8 mitochondrial genes.";
DR J. Anim. Sci. 76:2207-2208(1998).
KW EMBL; AF039170; AAD05063.1; -.
FT NON_TER
SQ SEQUENCE 15 AA; 1788 MW; 2B93E0B6A0588CB3 CRC64;

```

```

Query Match 22.9%; Score 25; DB 8; Length 15;
Best Local Similarity 40.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 2 RYFNOLSTGL 11
:|:|:|:|:|
DB 3 KYFEKWTSM 12

```

```

RESULT 11
0994F2 PRELIMINARY; PRT; 16 AA.
AC 0994F2:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RES-701-2-ENDOTHELIN TYPE B RECEPTOR ANTAGONIST.
OS Streptomyces.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae.
OX NCBI_TaxID=1883;
RN [1]
RP SEQUENCE.
RX MEDLINE=96139410; PubMed=8557586;
RA Yano K., Yamasaki M., Yoshida M., Matsuda Y., Yamaguchi K.;
RT "RES-701-2, a novel and selective endothelin type B receptor
RT antagonist produced by Streptomyces sp. II. Determination of the
RT primary structure.";
RL J. Antibiot. 48:1368-1370(1995).
SQ SEQUENCE 16 AA; 2061 MW; 4E7B0E0789E244DB CRC64;

```

```

Query Match 22.9%; Score 25; DB 2; Length 16;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 15 GLAADV 20
:|:|:|:|:|
DB 5 GTAPDW 10

```

```

RESULT 12
0995E9 PRELIMINARY; PRT; 16 AA.
ID 0995E9

```

```

AC 0995E9:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
OS Haemophilus somnus.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=731;
RN [1]
RP SEQUENCE.
RX MEDLINE=93114910; PubMed=8418069;
RA Tagawa Y., Ishikawa H., Yusa N.;
RT "Purification and partial characterization of the major outer membrane
RT protein of Haemophilus somnus.";
RL Infect. Immun. 61:91-96(1993).
SQ SEQUENCE 16 AA; 1709 MW; 910FCA6B728D0D65 CRC64;

```

```

Query Match 22.9%; Score 25; DB 2; Length 16;
Best Local Similarity 41.7%; Pred. No. 1.4e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 4 FNOJSTGLDMVG 15
:|:|:|:|:|
DB 4 YNONGTKVDVG 15

```

```

RESULT 13
090VJ9 PRELIMINARY; PRT; 19 AA.
ID 090VJ9
AC 090VJ9:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE MEPRIN-B PEPTIDE B1.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE.
RX MEDLINE=91373354; PubMed=1894622;
RA Kounnas M.Z., Wolz R.L., Gorbica C.M., Bond J.S.;
RT "Meprin-A and -B: Cell surface endopeptidases of the mouse kidney.";
RL J. Biol. Chem. 266:17350-17357(1991).
SQ SEQUENCE 19 AA; 2157 MW; 81E66F19417E20C5 CRC64;

```

```

Query Match 22.9%; Score 25; DB 11; Length 19;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 4 FNOJSTGLDMVG 15
:|:|:|:|:|
DB 1 FNOVSITNDNIG 12

```

```

RESULT 14
047529 PRELIMINARY; PRT; 22 AA.
ID 047529
AC 047529:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE ORF1 AND ORF2.
GN HEMA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-K-12; MEDLINE-93171869; PubMed-7679718;
 RA Post D.A., Hore-Jensen B., Switzer R.L.;
 RT "Characterization of the hemA-prs region of the Escherichia coli and
 RT Salmonella typhimurium chromosomes: Identification of two open reading
 RT frames and implications for prs expression";
 RL J. Gen. Microbiol. 139:259-266(1993).
 DR EMBL: M77237; AAA24432.1; -
 SQ SEQUENCE 22 AA; 2340 MW; 25802412BB5B4E51 CRC64;

Query Match 22.9%; Score 25; DB 2; Length 22;
 Best Local Similarity 38.5%; Pred. No. 2e+03;
 Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 STGLDVGGLADW 20
 : : : : :
 Db 6 NTNDVRSLLASY 18

RESULT 15
 ID 048661 PRELIMINARY; PRT; 22 AA.
 AC 048661;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
 DE HOMEBOX PROTEIN (FRAGMENT).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
 NC NCB1_TaxID-4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IRB21;
 RA Liu G., Yang J., Zhai W., He P., Li X., Lu J., Li S., Zhu L.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF003600; AAB94485.1; -
 KW Homebox; DNA-binding; Nuclear protein.
 FT NON_TER 1
 FT NON_TER 22
 SQ SEQUENCE 22 AA; 2337 MW; B18E95E52F6CE9C CRC64;

Query Match 22.9%; Score 25; DB 10; Length 22;
 Best Local Similarity 100.0%; Pred. No. 2e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STGLD 12
 : : : : :
 Db 8 STGLD 12

RESULT 16
 ID 09T2J3 PRELIMINARY; PRT; 23 AA.
 AC 09T2J3;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE 8 KDA PHOTOSYSTEM I PSBG PROTEIN (FRAGMENT).
 OS Nicotiana glauca (Wood tobacco).
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Nicotiana.
 NC NCB1_TaxID-4096;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-94105345; PubMed-8278548;
 RA Oookata J., Mikami K., Hayashida N., Nakamura M., Sugiyama M.;
 RT "Molecular heterogeneity of photosystem I. psad, psae, psaf, psah, and
 RT psal are all present in isoforms in Nicotiana spp.";

RL Plant Physiol. 102:1259-1267(1993).
 SQ SEQUENCE 23 AA; 2415 MW; E56CC139D095A968 CRC64;

Query Match 22.9%; Score 25; DB 8; Length 23;
 Best Local Similarity 71.4%; Pred. No. 2.1e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 LSTGLDM 13
 : : : : :
 Db 10 LSTGLSL 16

RESULT 17
 ID 09QV05 PRELIMINARY; PRT; 16 AA.
 AC 09QV05;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE IRI HEAVY CHAIN-RELATED PROTEIN (FRAGMENT).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID-10118;
 RN [1]
 RP SEQUENCE.
 RA Choi-Miura N.H., Sano Y., Oda E., Nakano Y., Tobe T., Yanagishita T.,
 RA Tanigawa M., Katagiri T., Tomita M.;
 RA J. Biochem. 117:400-407(1995).
 SQ SEQUENCE 16 AA; 1809 MW; 89436CB4C81609C7 CRC64;

Query Match 22.0%; Score 24; DB 11; Length 16;
 Best Local Similarity 40.0%; Pred. No. 2e+03;
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 10 GLDMVGLAD 19
 : : : : :
 Db 4 GIDIVSLTVD 13

RESULT 18
 ID 069392 PRELIMINARY; PRT; 23 AA.
 AC 069392;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE THE FIRST 23 CODONS OF PRV GC ALLELE ASGRL PRECURSOR (FRAGMENT).
 OS GC.
 OS Pseudorabies virus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicelloviruses.
 NC NCB1_TaxID-10345;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BECKER;
 RX MEDLINE-95081163; PubMed-7989378;
 RA Tomlino M., Wilkinson K.S., Ryan P.;
 RT "Can a signal sequence become too hydrophobic?";
 RL J. Biol. Chem. 269:32016-32021(1994).
 DR EMBL: L36971; AAA79968.1; -
 KW Signal.
 FT SIGNAL 1
 FT SIGNAL 22
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2188 MW; B7069865A1406F3F CRC64;

Query Match 22.0%; Score 24; DB 13; Length 23;
 Best Local Similarity 54.5%; Pred. No. 3.1e+03;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 8 STGLDMVGLAA 18
1 1 1 1 1
DB 3 SIGLAMLALLA 13

RESULT 19
O9S941 PRELIMINARY; PRT; 13 AA.

AC 09S941: PRELIMINARY; PRT; 13 AA.
ID 09S941: PRELIMINARY; PRT; 13 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE H(+)-TRANSLOCATING (PYROPHOSPHATE-ENERGIZED) INORGANIC
DE PYROPHOSPHATASE BETA-2 POLYPEPTIDE (EC 3.6.1.1) (FRAGMENT).
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons, core eudicots; Caryophyllales;
OC Caryophyllales; Chenopodiaceae; Beta.
OX NCBI_TaxID=3555;
RN [1]
RP MEDLINE=92179265; PubMed=1311852;
RA Sarafian V., Kim Y., Poole R.J., Rea P.A.;
RT "Molecular cloning and sequence of cDNA encoding the pyrophosphate-
energized vacuolar membrane proton pump of Arabidopsis thaliana.";
RT Proc. Natl. Acad. Sci. U.S.A. 89:1775-1779(1992).
SQ SEQUENCE 13 AA; 1243 MW; C9DB193C37C22AD CRC64;

Query Match 21.1%; Score 23; DB 10; Length 13;
Best Local Similarity 66.7%; Pred. No. 2.3e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 10 GLDMVG 15
1 1 1 1 1
DB 5 GADLVG 10

RESULT 20
O9PXC5 PRELIMINARY; PRT; 15 AA.

AC 09PXC5: PRELIMINARY; PRT; 15 AA.
ID 09PXC5: PRELIMINARY; PRT; 15 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE SMALL NUCLEAR INCLUSION POLYPEPTIDE CLEAVAGE PRODUCT.
OS Tobacco etch virus (TEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
OC Potyvirus.
OX NCBI_TaxID=12227;
RN [1]
RP MEDLINE=91306428; PubMed=1853555;
RA Dougherty W.G., Parks T.D.;
RT "Post-translational processing of the tobacco etch virus 49-kDa small
nuclear inclusion polypeptide: identification of an internal cleavage
site and delimitation of VPg and proteinase domains.";
RT Virology 183:449-456(1991).
SQ SEQUENCE 15 AA; 1680 MW; 2273B6E461D0F28C CRC64;

Query Match 21.1%; Score 23; DB 12; Length 15;
Best Local Similarity 50.0%; Pred. No. 2.8e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 PRYFNOLS 8
1 1 1 1 1
DB 8 PRDYNPIS 15

RESULT 21
O9UCP7 PRELIMINARY; PRT; 17 AA.

AC 09UCP7: PRELIMINARY; PRT; 17 AA.
ID 09UCP7: PRELIMINARY; PRT; 17 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HML-1 BETA SUBUNIT, SMALLER HML-1 SUBUNIT, BETA 7
DE INTEGRIN-HETERODIMERIC PROTEIN COMPLEX RECOGNIZED BY THE HUMAN
DE MUCOSAL LYMPHOCYTE 1 MONOCLONAL ANTIBODY.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=92179294; PubMed=1542691;
RA Parker C.M., Ceppek K.L., Russell G.J., Shaw S.K., Posnett D.N.,
RA Schvarling R., Brenner M.B.;
RT "A family of beta 7 integrins on human mucosal lymphocytes.";
RT Proc. Natl. Acad. Sci. U.S.A. 89:1924-1928(1992).
SQ SEQUENCE 17 AA; 1903 MW; 053C931745A8BA2 CRC64;

Query Match 21.1%; Score 23; DB 4; Length 17;
Best Local Similarity 26.7%; Pred. No. 3.2e+03;
Matches 4; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 6 QUSTGLDMVGLADW 20
1 1 1 1 1 1 1 1 1 1
DB 1 ELDAKIPSTGDATFW 15

RESULT 22
O9TRD9 PRELIMINARY; PRT; 17 AA.

AC 09TRD9: PRELIMINARY; PRT; 17 AA.
ID 09TRD9: PRELIMINARY; PRT; 17 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE V2 VASOPRESSIN RECEPTOR (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP MEDLINE=94079886; PubMed=8257689;
RA Kojima E., Eich P., Gimpl G., Fahrenholz F.;
RT "Direct identification of an extracellular agonist binding site in the
renal V2 vasopressin receptor.";
RT Biochemistry 32:13537-13544(1993).
SQ SEQUENCE 17 AA; 1904 MW; E3C89EB346200E57 CRC64;

Query Match 21.1%; Score 23; DB 6; Length 17;
Best Local Similarity 62.5%; Pred. No. 3.2e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 11 LDMVGLAA 18
1 1 1 1 1 1 1 1 1 1
DB 9 LQWVGMYA 16

RESULT 23
O24445 PRELIMINARY; PRT; 17 AA.

AC 024445: PRELIMINARY; PRT; 17 AA.
ID 024445: PRELIMINARY; PRT; 17 AA.
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
DE THIS ORF RESIDES IN THE 5' UTR OF CPHY2.
OS Ceratodon purpureus (Moss).
OC Eukaryota; Viridiplantae; Embryophyta; Bryophyta; Bryopsida;
OC Dicranidae; Dicranales; Ditrichaceae; Ceratodon.

OX NCBI_TaxID-3225;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WT3;
 RA Hughes J.E., Lamparter T., Mittmann F.;
 RL Plant Physiol. 112:446-446(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WT3;
 RA Presentis K., Paulo N., Dittlich P., Algarra P., Thuenmler F.;
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U72993; AAB19059.1;
 SQ SEQUENCE 17 AA; 1949 MW; 072DF72059DF1C7C CRC64;

Query Match 21.1%; Score 23; DB 10; Length 17;
 Best Local Similarity 38.5%; Pred. NO. 3.2e+03;
 Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 4 FNOLSTGLDWGL 16
 | : | : | :
 Db 4 FSTRSRSLMIVGI 16

RESULT 24
 Q9S888
 ID Q9S888 PRELIMINARY; PRT; 18 AA.
 AC Q9S888;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE CHAPERONIN 60 BETA CHAIN (FRAGMENT).
 OS Narcissus pseudonarcissus (Daifodil).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae; Narcissus.
 OX NCBI_TaxID-39639;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-96291727; PubMed-8754688;
 RA Bonk M., Tardos M., Vandekerckhove J., Al-Babili S., Beyer P.;
 RT "Purification and characterization of chaperonin 60 and heat-shock
 protein 70 from chromoplasts of Narcissus pseudonarcissus.";
 RL Plant Physiol. 111:931-939(1996).
 SQ SEQUENCE 18 AA; 2108 MW; E324CCA84907EE8 CRC64;

Query Match 21.1%; Score 23; DB 10; Length 18;
 Best Local Similarity 44.4%; Pred. NO. 3.4e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 YFNOLSTGL 11
 ||| : | :
 Db 5 YFNKDGTAI 13

RESULT 25
 Q9T2J2
 ID Q9T2J2 PRELIMINARY; PRT; 19 AA.
 AC Q9T2J2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE 8 KDA PHOTOSYSTEM I PSAG PROTEIN (FRAGMENT).
 OS Nicotiana tabacum (Common tobacco).
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID-4097;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-94105345; PubMed-8278548;
 RA Obokata J., Mikami K., Hayashida N., Nakamura M., Sugitara M.;

RT "Molecular heterogeneity of photosystem I. psad, psae, psaf, psah, and
 RT psal are all present in isoforms in Nicotiana spp.";
 RL Plant Physiol. 102:1259-1267(1993).
 SQ SEQUENCE 19 AA; 1976 MW; E08869791A6712FE CRC64;

Query Match 21.1%; Score 23; DB 8; Length 19;
 Best Local Similarity 100.0%; Pred. NO. 3.6e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LSTGL 11
 |||||
 Db 10 LSTGL 14

Search completed: February 5, 2001, 10:55:27
 Job time: 916 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:47:33 ; Search time 68.99 Seconds
(without alignments)
9.913 Million cell updates/sec

Title: US-08-981-824-5

Perfect score: 109
Sequence: 1 PRFNOGLSTGLDMVGLADW 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 segs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_36:*

1:	/SIDSL/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2:	/SIDSL/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3:	/SIDSL/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4:	/SIDSL/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5:	/SIDSL/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6:	/SIDSL/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7:	/SIDSL/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8:	/SIDSL/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9:	/SIDSL/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10:	/SIDSL/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11:	/SIDSL/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12:	/SIDSL/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13:	/SIDSL/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14:	/SIDSL/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15:	/SIDSL/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16:	/SIDSL/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17:	/SIDSL/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18:	/SIDSL/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19:	/SIDSL/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20:	/SIDSL/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21:	/SIDSL/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	100.0	20	18	W18846
2	109	100.0	20	18	W01197
3	84	77.1	20	16	W72272
4	84	77.1	20	21	Y59572
5	77	70.6	15	18	W12405
6	73	67.0	14	16	R76657
7	73	67.0	14	18	W18865
8	65	59.6	15	18	W12404
9	55	50.5	20	18	W18845
10	55	50.5	20	18	W01196
11	50	45.9	20	16	R72273
12	50	45.9	20	21	Y59573

13	36	33.0	14	19	W70362
14	36	33.0	14	19	W70363
15	36	33.0	15	19	W70361
16	36	33.0	22	16	R65046
17	36	33.0	22	19	W46649
18	36	33.0	22	20	Y29339
19	32	29.4	15	21	Y79295
20	32	29.4	15	21	Y67045
21	32	29.4	15	21	Y54682
22	31	28.4	18	16	W13198
23	31	28.4	18	19	W57351
24	31	28.4	19	21	Y79452
25	31	28.4	20	21	Y19366
26	31	28.4	24	16	W13206
27	31	28.4	24	19	W57367
28	30	27.5	17	20	Y27549
29	30	27.5	18	16	R66373
30	30	27.5	18	16	R66374
31	30	27.5	19	15	R65378
32	30	27.5	19	16	R82589
33	30	27.5	19	17	W05615
34	30	27.5	19	21	Y80074
35	30	27.5	19	21	Y54556
36	30	27.5	19	21	Y58780
37	30	27.5	20	18	W33044
38	30	27.5	20	18	W12352
39	30	27.5	21	13	R22911
40	29	26.6	9	20	Y01008
41	29	26.6	12	18	W29922
42	29	26.6	13	16	R64138
43	29	26.6	13	17	R97212
44	29	26.6	13	19	W43316
45	29	26.6	14	16	R76647

ALIGNMENTS

RESULT 1	
W18846	W18846 standard; peptide; 20 AA.
XX	XX
AC	W18846;
XX	XX
DT	05-JAN-1998 (first entry)
XX	XX
DE	65 kD Glutamic acid decarboxylase peptide fragment V.
XX	XX
KW	GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
KW	Insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
KW	predisposition; autoimmune; tumour; rheumatoid arthritis;
KW	multiple sclerosis.
XX	XX
OS	Synthetic.
XX	XX
FN	DE19526561-AL.
XX	XX
PD	23-JAN-1997.
XX	XX
PF	20-JUL-1995; 95DE-1026561.
XX	XX
PR	20-JUL-1995; 95DE-1026561.
XX	XX
PA	(BOEF) BOEHRINGER MANNHEIM GMBH.
XX	XX
PI	Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;
PI	Pozzilli P, Stahl P;
XX	XX
DR	WPI; 1997-088254/09.
XX	XX
PT	skin test for diagnosis of cell-mediated diseases, esp. diabetes -
XX	involving intradermal admn. of auto-reactive substances

PS Claim 11; Page 9; 12pp; German.
 XX
 CC W18842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predisposition to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 109; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.2e-13;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRYFNOLSTGLDMVGLADW 20
 |||||
 Db 1 pryfnglstgldmvglaadw 20

RESULT 2

W01797
 ID W01797 standard; peptide: 20 AA.

XX
 AC W01797;

XX
 DT 15-OCT-1997 (first entry)

XX
 DE Human 65 kD glutamine decarboxylase peptide.

XX
 KM Human; glutamine decarboxylase; GAD; diagnosis; predisposition;
 KM tumour; immunological; disease; autoimmune; diabetes; reagent;
 KM determination; T cell; subpopulation; medication; treatment;
 KM prevention; production; antigen; immunogen; tolerogen; isolation;
 KM reinfection; inactivation.
 XX

OS Homo sapiens.

XX
 PN DE19525784-A1.

XX
 PD 16-JAN-1997.

XX
 PF 14-JUL-1995; 95DE-1025784.

XX
 PR 14-JUL-1995; 95DE-1025784.

XX
 PA (BOE) BOEHRINGER MANNHEIM GMBH.

XX
 PI Albert W, Boltard C, Endl J, Jung G, Schendel D;
 PI Stahl P, Van Endert P;

XX
 DR WPI: 1997-078452/08.

XX
 PT Glutamine decarboxylase peptide(s) - for diagnosis and therapy of
 PT diabetes, etc.
 XX

PS Claim 1; Page 12; 15pp; German.

XX
 CC The present peptide is a fragment of the human 65 kD glutamine
 CC decarboxylase (GAD), which can be used to diagnose, or diagnose a
 CC predisposition to, a tumour or immunological disease, preferably an
 CC autoimmune disease, especially diabetes. It can also be used as a
 CC reagent to determine specific T cell subpopulations, in medicaments
 CC to treat or prevent immunological diseases, preferably autoimmune
 CC diseases, especially diabetes, to produce antigens, especially
 CC immunogens or tolerogens and to isolate specific T cell
 CC subpopulations, which can be used to produce antigens or for
 CC reinfection, optionally after inactivation.

XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 109; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.2e-13;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRYFNOLSTGLDMVGLADW 20
 |||||
 Db 1 pryfnglstgldmvglaadw 20

RESULT 3

R72272
 ID R72272 standard; Peptide: 20 AA.

XX
 AC R72272;

XX
 DT 13-NOV-1995 (first entry)

XX
 DE Glutamic acid decarboxylase (GAD65) fragment.

XX
 KM Glutamic acid decarboxylase; GAD65; autoimmune disorders;
 KM insulin-dependent diabetes mellitus; stiff man disease.
 XX

XX
 OS Homo sapiens.

XX
 PN W09507992-A.

XX
 PD 23-MAR-1995.

XX
 PF 24-AUG-1994; 94WO-US09478.

XX
 PR 17-SEP-1993; 93US-0123859.

XX
 PA (REGC) UNIV CALIFORNIA.

XX
 PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;

XX
 DR WPI: 1995-131360/17.

XX
 PT New polypeptide fragments of glutamic acid decarboxylase - for
 PT diagnosis and treatment of autoimmune disease, esp. insulin
 PT dependent diabetes, also related nucleic acid, vectors,
 PT antibodies, hybridoma(s) etc.
 XX

PS Example 11; Page 76; 100pp; English.

XX
 CC Q86481 and Q86482 encode R7173 and R79105, rat and human glutamic
 CC acid decarboxylase (GAD65) respectively, from which the GAD65
 CC fragments described in R72261-R72298 were derived. These fragments
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose
 CC and treat GAD-related autoimmune disorders, such as insulin
 CC dependent diabetes mellitus or stiff man disease.
 XX

XX
 SQ Sequence 20 AA;

Query Match 77.1%; Score 84; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.3e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRYFNOLSTGLDMVGL 16
 |||||
 Db 5 pryfnglstgldmvgyl 20

RESULT 4

Y59572
 ID Y59572 standard; peptide: 20 AA.

XX
 AC Y59572;

```

XX 03-APR-2000 (first entry)
DT GAD65 fragment, peptide #12.
XX
DE GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
KW insulin dependent diabetes mellitus; stiff man disease; diagnosis;
KM therapy.
XX
OS Homo sapiens.
XX
PN US5998366-A.
XX
PD 07-DEC-1999.
XX
PF 09-APR-1997; 97US-0827618.
XX
PR 07-JUN-1995; 95US-0485725.
PR 21-SEP-1990; 90US-0586536.
PR 18-JUN-1991; 91US-0716909.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Tobin AJ, Kaufman DL, Erlander MG;
DR WPI; 2000-095930/08.
XX
PT Ameliorating glutamic acid decarboxylase associated autoimmune
PT disorders such as insulin dependent diabetes mellitus and stiff man
PT disease -
XX
PS Example 11; Column 42; 61pp; English.
XX
CC This sequence represents a fragment of the glutamic acid decarboxylase 65
CC (GAD65) protein. The invention relates to a method of ameliorating GAD
CC associated autoimmune disorders by administering a GAD65 peptide to the
CC patient. The method can be used for ameliorating GAD associated
CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
CC and stiff man disease. GAD65 can also be useful for screening drugs that
CC alter GAD function, for generating monoclonal antibodies and in
CC immunosays. GAD65 is an effective diagnostic tool for predicting IDDM
CC and the diagnosis is quite easy. It is also possible to obtain much
CC larger quantities of polypeptide via recombinant techniques than are
CC available from natural sources.
XX
SQ Sequence 20 AA:
XX
Query Match 77.18; Score 84; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 PRYFNQSTGLDMVGL 16
Db 5 prfngqstgldmvgl 20
XX
RESULT 5
W12405
ID W12405 standard; peptide; 15 AA.
XX
AC W12405;
XX
DE 08-OCT-1997 (first entry)
DE GAD65 residues 177-191.
XX
KW 65 kD glutamic acid decarboxylase; human; GAD65; enzyme; pancreas;
KW neuron; central nervous system; type I diabetes; autoimmune response;
KW T cell; therapy.
XX
OS Homo sapiens.
XX

```

```

FH Key Location/Qualifiers
FT MISC-difference 1..15
FT FT /note- "optionally substituted, providing at least 7
FT FT residues remain wild type"
XX
XX W09700891-A1.
XX
PD 09-JAN-1997.
XX
PF 24-JUN-1996; 96MO-US10790.
XX
PR 23-JUN-1995; 95US-0494624.
XX
PA (KENN-) KENNEDY INST RHEUMATOLOGY.
PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
PA (SBR-) ST BARTHOLOMEW'S HOSPITAL CENT CLINICAL.
XX
PI Conlon PJ, Gaur A, Leslie RDG, Ling N, Londel M;
DR WPI; 1997-087322/08.
XX
PT New human glutamic acid decarboxylase peptide(s) - used for
PT treatment, diagnosis and determining predisposition to diabetes and
PT for ameliorating auto-immune responses.
XX
PS Claim 23; Page -: 28pp; English.
XX
CC W12403-W12413 represent fragments and analogues of the the 65 kD isoform
CC of human glutamic acid decarboxylase (GAD65) (see W12402 for full length
CC wild type protein). GAD is an enzyme expressed in the beta cells of the
CC pancreas, and in neurons of the central nervous system. There are two
CC isoforms of GAD, a 67 kD isoform, and GAD65. Immunodominant regions of
CC GAD65 have been identified in type I diabetic patients. These GAD65
CC fragments, and analogues, are used in the methods of the invention. The
CC methods are for detecting or treating diabetes or a predisposition to
CC diabetes. The peptides can also be used for ameliorating an autoimmune
CC response in a patient. Alteration of the native peptides with selective
CC changes of crucial residues can induce unresponsiveness or change the
CC responsiveness of antigen-specific autoreactive T cells. The peptide
CC analogues compete for binding to MHC and do not cause proliferation of
CC the corresponding native peptide-specific T cells.
XX
SQ Sequence 15 AA:
XX
Query Match 70.6%; Score 77; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 RYFNQSTGLDMVGL 16
Db 1 rfyngqstgldmvgl 15
XX
RESULT 6
R76657
ID R76657 standard; peptide; 14 AA.
XX
AC R76657;
XX
DE 05-MAR-1996 (first entry)
DE Peptide derived from human glutamic acid decarboxylase 16.
XX
KW diabetes; T-cell subpopulation; detection; antigen production;
KW diagnosis; autoimmune disease.
XX
OS Homo sapiens.
XX
PN DE4418091-A1.
XX
PD 27-JUL-1995.
XX

```

PF 24-MAY-1994; 94DE-4418091.
 XX
 XX 04-FEB-1994; 94DE-4403522.
 PR 20-JAN-1994; 94DE-4401629.
 XX
 PA (ENDL/) ENDL J.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Albert W, Dormair K, Endl J, Jung G, Meinel E;
 PI Stahl P, Schendel D;
 XX WPI: 1995-264505/35.
 DR
 XX
 PT Antigen-specific activated T-lymphocytes and their detection - by
 PT interaction with inventive peptide(s) of peptide-MHC complexes;
 PT useful in diagnosis of e.g. diabetes and auto-immune diseases
 XX
 PS Claim 1, Fig 2; 21pp; German.
 XX
 CC R76642-62 are derived from human glutamic acid decarboxylase and
 CC specifically react with T-cell sub-populations isolated from recently
 CC diagnosed Type-I diabetics. Pharmaceutical compns. contg. these
 CC peptides and those shown in R77571-72, are useful for the diagnosis of
 CC a disease or predispositions of immune system diseases, tumours, and
 CC autoimmune diseases, including diabetes. The peptides are able to detect
 CC specific T-cell subpopulations that are then used for antigen prodn.,
 CC e.g. by reinjection.
 CC
 SQ Sequence 14 AA;

Query Match 67.0%; Score 73; DB 16; Length 14;
 Best Local Similarity 100.0%; Pred. No. 8.2e-07;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYFNOLSTGLDMVG 15
 |||||
 Db 1 RYFNGLSTGLDMVG 14

RESULT 7
 W18865
 ID W18865 standard; peptide: 14 AA.
 XX
 AC W18865;
 XX
 DT 05-JAN-1998 (first entry)
 XX
 DE 65 KD Glutamic acid decarboxylase peptide fragment 16.
 XX
 KW GAD; 65 KD: human; glutamic acid decarboxylase; autoreactive; diagnosis;
 KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
 KW predilection; autoimmune; tumour; rheumatoid arthritis;
 KW multiple sclerosis.
 XX
 OS Synthetic.
 XX
 PN DE19526561-A1.
 XX
 PD 23-JAN-1997.
 XX
 PF 20-JUL-1995; 95DE-1026561.
 XX
 PR 20-JUL-1995; 95DE-1026561.
 XX
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;
 PI Pozzilli P, Stahl P;
 XX
 DR WPI: 1997-088254/09.
 XX
 PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -

PF Involving intradermal admin. of auto-reactive substances
 XX
 PS Claim 11; Fig 2; 12pp; German.
 XX
 CC W18842-70 are peptide fragments of the 65 KD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predilection to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 CC
 SQ Sequence 14 AA;

Query Match 67.0%; Score 73; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 8.2e-07;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYFNOLSTGLDMVG 15
 |||||
 Db 1 RYFNGLSTGLDMVG 14

RESULT 8
 W12404
 ID W12404 standard; peptide: 15 AA.
 XX
 AC W12404;
 XX
 DT 08-OCT-1997 (first entry)
 XX
 DE GAD65 residues 173-187.
 XX
 KW 65 KD glutamic acid decarboxylase: human; GAD65; enzyme; pancreas;
 KW neuron; central nervous system; type I diabetes; autoimmune response;
 KW T cell; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1.15 /note= "optionally substituted, providing at least 7
 FT residues remain wild type"

W09700891-A1.
 PD 09-JAN-1997.
 XX
 PF 24-JUN-1996; 96WO-US10790.
 XX
 PR 23-JUN-1995; 95US-0494624.
 XX
 PA (KENN-) KENNEDY INST RHEUMATOLOGY.
 PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
 PA (SBAR-) ST BARTHOLOMEW'S HOSPITAL CENT CLINICAL.
 PI Conlon PJ, Gaur A, Leslie RDG, Ling N, Londei M;
 DR WPI: 1997-087322/08.
 XX
 PT New human glutamic acid decarboxylase peptide(s) - used for
 PT treatment, diagnosis and determining predilection to diabetes and
 PT for ameliorating auto-immune responses.
 XX
 PS Claim 22; Page -; 28pp; English.
 XX
 CC W12403-W12413 represent fragments and analogues of the 65 KD isoform
 CC of human glutamic acid decarboxylase (GAD65) (see W12402 for full length
 CC wild type protein). GAD is an enzyme expressed in the beta cells of the

CC pancreas, and in neurons of the central nervous system. There are two
 CC isoforms of GAD, a 67 kD isoform, and GAD65. Immunodominant regions of
 CC GAD65 have been identified in Type I diabetic patients. These GAD65
 CC fragments, and analogues, are used in the methods of the invention. The
 CC methods are for detecting or treating diabetes or a predisposition to
 CC diabetes. The peptides can also be used for ameliorating an autoimmune
 CC response in a patient. Alteration of the native peptides with selective
 CC changes of crucial residues can induce unresponsiveness or change the
 CC responsiveness of antigen-specific autoreactive T cells. The peptide
 CC analogues compete for binding to MHC and do not cause proliferation of
 CC the corresponding native peptide-specific T cells.

XX Sequence 15 AA:

Query Match 59.6%; Score 65; DB 18; Length 15;
 Best Local Similarity 100.0%; Pred. No. 2.5e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRYFNQSTGSD 12
 |||||
 Db 4 pryfngstgld 15

RESULT 9
 W18845
 ID W18845 standard; peptide: 20 AA.
 XX W18845;
 AC
 XX
 DT 05-JAN-1998 (first entry)
 XX
 DE 65 kD glutamic acid decarboxylase peptide fragment IV.
 XX
 GAD: 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
 KM insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
 KM predisposition; autoimmune; tumour; rheumatoid arthritis;
 KM multiple sclerosis.
 XX
 OS Synthetic.
 XX
 PN DE19526561-A1.
 XX
 PD 23-JAN-1997.
 XX
 PF 20-JUL-1995; 95DE-1026561.
 XX
 PR 20-JUL-1995; 95DE-1026561.
 XX
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 XX
 PI Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;
 PI Pozzilli P, Stahl P;
 XX
 DR WPI; 1997-088254/09.
 XX
 PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
 PT Involving intradermal admin. of auto-reactive substances
 XX
 PS Claim 11; Page 9; 12pp; German.
 XX
 CC W18842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predisposition to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 XX
 SQ Sequence 20 AA:

Query Match 50.5%; Score 55; DB 18; Length 20;
 * Best Local Similarity 100.0%; Pred. No. 0.0022;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRYFNQSTG 10
 |||||
 Db 11 pryfngstg 20

RESULT 10

W01796
 ID W01796 standard; peptide: 20 AA.

XX W01796;
 AC
 XX
 DT 15-OCT-1997 (first entry)
 XX
 DE Human 65 kD glutamine decarboxylase peptide.

XX
 KM Human; glutamine decarboxylase; GAD; diagnosis; predisposition;
 KM tumour; immunological; disease; autoimmune; diabetes; reagent;
 KM determination; T cell; subpopulation; medicament; treatment;
 KM prevention; production; antigen; immunogen; tolerogen; isolation;
 KM reinjection; inactivation.
 XX
 OS Homo sapiens.
 XX
 PN DE19525784-A1.
 XX
 PD 16-JAN-1997.
 XX
 PF 14-JUL-1995; 95DE-1025784.
 XX
 PR 14-JUL-1995; 95DE-1025784.
 XX
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.

XX
 PI Albert W, Boltard C, Endl J, Jung G, Schendel D;
 PI Stahl P, Van Endert P;
 XX
 DR WPI; 1997-078452/08.
 XX
 PT Glutamine decarboxylase peptide(s) - for diagnosis and therapy of
 PT diabetes, etc.
 XX
 PS Claim 1; Page 12; 15pp; German.

XX
 CC The present peptide is a fragment of the human 65 kD glutamine
 CC decarboxylase (GAD), which can be used to diagnose, or diagnose a
 CC predisposition to, a tumour or immunological disease, preferably an
 CC autoimmune disease, especially diabetes. It can also be used as a
 CC reagent to determine specific T cell subpopulations, in medicaments
 CC to treat or prevent immunological diseases, preferably autoimmune
 CC diseases, especially diabetes, to produce antigens, especially
 CC immunogens or tolerogens and to isolate specific T cell
 CC subpopulations, which can be used to produce antigens or for
 CC reinjection, optionally after inactivation.
 XX
 SQ Sequence 20 AA:

Query Match 50.5%; Score 55; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.0022;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRYFNQSTG 10
 |||||
 Db 11 pryfngstg 20

RESULT 11

```

R72273
ID R72273 standard; Peptide; 20 AA.
XX
AC R72273;
XX
DE 13-NOV-1995 (first entry)
XX
DE Glutamic acid decarboxylase (GAD65) fragment.
XX
KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
XX insulin-dependant diabetes mellitus; stiff man disease.
XX
OS Homo sapiens.
XX
PN W09507992-A.
XX
PD 23-MAR-1995.
XX
PF 24-AUG-1994; 94MO-US09478.
XX
PR 17-SEP-1993; 93US-0123859.
XX
PA (RESC ) UNIV CALIFORNIA.
XX
PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
XX WPI: 1995-131360/17.
XX
DR New polypeptide fragments of glutamic acid decarboxylase - for
XX diagnosis and treatment of autoimmune disease, esp. insulin
XX dependent diabetes, also related nucleic acid, vectors,
XX antibodies, hydridoma(s) etc.
XX
PS Example 11; Page 76; 100pp; English.
XX
CC 086481 and 086482 encode R71733 and R79105, rat and human glutamic
XX acid decarboxylase (GAD65) respectively, from which the GAD65
XX fragments described in R72261-R72298 were derived. These fragments
XX can be used to detect autoantibodies against GAD, e.g. to diagnose
XX CC and treat GAD-related autoimmune disorders, such as insulin
XX dependent diabetes mellitus or stiff man disease.
XX
SQ Sequence 20 AA;

Query Match 45.9%; Score 50; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DMVGLAADM 20
   |||||
   1 dmvglaadm 9

Db

RESULT 12
V59573
ID Y59573 standard; peptide; 20 AA.
XX
AC Y59573;
XX
DE 03-APR-2000 (first entry)
XX
DE GAD65 fragment, peptide #13.
XX
KW GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
XX insulin, dependent diabetes mellitus; Stiff man disease; diagnosis;
XX therapy.
XX
OS Homo sapiens.
XX
PN US5998366-A;
XX
PD 07-DEC-1999.

```

```

XX
PF 09-APR-1997; 97US-0827618.
XX
PR 07-JUN-1995; 95US-0485725.
XX
PR 21-SEP-1990; 90US-0586536.
XX
PR 18-JUN-1991; 91US-0716909.
XX
PA (RESC ) UNIV CALIFORNIA.
XX
PI Tobin AJ, Kaufman DL, Erlander MG;
XX WPI: 2000-095930/08.
XX
DR WPI: 2000-095930/08.
XX
PF Ameliorating glutamic acid decarboxylase associated autoimmune
XX disorders such as insulin dependent diabetes mellitus and Stiff man
XX disease -
XX
PS Example 11; Column 42; 61pp; English.
XX
CC This sequence represents a fragment of the glutamic acid decarboxylase 65
XX (GAD65) protein. The invention relates to a method of ameliorating GAD
XX associated autoimmune disorders by administering a GAD65 peptide to the
XX patient. The method can be used for ameliorating GAD associated
XX autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
XX and Stiff man disease. GAD65 can also be useful for screening drugs that
XX alter GAD function, for generating monoclonal antibodies and in
XX immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
XX and the diagnosis is quite easy. It is also possible to obtain much
XX larger quantities of polypeptide via recombinant techniques than are
XX available from natural sources.
XX
SQ Sequence 20 AA;

```

```

Query Match 45.9%; Score 50; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DMVGLAADM 20
   |||||
   1 dmvglaadm 9

Db

RESULT 13
W70362
ID W70362 standard; Protein; 14 AA.
XX
AC W70362;
XX
DE 14-DEC-1998 (first entry)
XX
DE Epitope tag WTGLD, open reading frame 2.
XX
KW Target nucleotide; epitope; gene tagging; epitope tagging.
XX
OS Synthetic.
XX
PN W09826094-A1.
XX
PD 18-JUN-1998.
XX
PF 09-DEC-1997; 97MO-US22472.
XX
PR 09-DEC-1996; 96US-0762106.
XX
PA (JARV/) JARVIK J W.
XX
PI Jarvik JW;
XX
DR WPI: 1998-348546/30.
XX
DR N-PSDB; V33280.
XX
XX Epitope tagging of genes, transcripts and polypeptides - using an

```

PT oligonucleotide comprising a nucleotide sequence encoding the
PT epitope independently of the reading frame of the nucleotide
PT sequence

XX Disclosure: Page 15; 46pp; English.

CC The epitope tag is encoded by the nucleotide sequence independently of
CC the reading frame. The epitope and methods disclosed can be used for the
CC tagging of genes, transcripts and polypeptides. They can be used for
CC e.g. discovering new genes, determining the size and abundance of
CC proteins produced by newly discovered genes, tracking the movement of
CC proteins within cell membranes, monitoring receptor binding and
CC internalisation of exogenous proteins, identifying the components of
CC functional protein complexes, purifying proteins, discovering the
CC function of proteins. The products, methods and uses are particularly
CC used for proteins that are unstable, are difficult to purify, or share
CC epitopes with a number of other proteins. The methods overcome the
CC inefficiency of epitope tagging caused by reading frame obstacles and
CC orientation obstacles.

CC Sequence 14 AA;

Query Match 33.0%; Score 36; DB 19; Length 14;

Best Local Similarity 66.7%; Pred. No. 3.7;

Matches 8; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

OY 9 TGLDMVGLADW 20

Db 5 tglDMVGL--dW 14

RESULT 14

ID W70363 standard; Protein: 14 AA.

XX W70363;

DT 14-DEC-1998 (first entry)

DE Epitope tag WTGLD, open reading frame 3.

XX Target nucleotide: epitope; gene tagging; epitope tagging.

OS Synthetic.

PN WO9826094-A1.

XX 18-JUN-1998.

PF 09-DEC-1997; 97WO-US22472.

PR 09-DEC-1996; 96US-0762106.

PA (JARV/) JARVIK J W.

XX Jarvik JW;

DR WPI: 1998-348546/30.

XX N-PSDB; V33280.

PT Epitope tagging of genes, transcripts and polypeptides - using an
PT oligonucleotide comprising a nucleotide sequence encoding the
PT epitope independently of the reading frame of the nucleotide
PT sequence

PS Disclosure: Page 15; 46pp; English.

CC The epitope tag is encoded by the nucleotide sequence independently of
CC the reading frame. The epitope and methods disclosed can be used for the
CC tagging of genes, transcripts and polypeptides. They can be used for
CC e.g. discovering new genes, determining the size and abundance of
CC proteins produced by newly discovered genes, tracking the movement of

CC proteins within cell membranes, monitoring receptor binding and
CC internalisation of exogenous proteins, identifying the components of
CC functional protein complexes, purifying proteins, discovering the
CC function of proteins. The products, methods and uses are particularly
CC used for proteins that are unstable, are difficult to purify, or share
CC epitopes with a number of other proteins. The methods overcome the
CC inefficiency of epitope tagging caused by reading frame obstacles and
CC orientation obstacles.

CC Sequence 14 AA;

Query Match 33.0%; Score 36; DB 19; Length 14;

Best Local Similarity 66.7%; Pred. No. 3.7;

Matches 8; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

OY 9 TGLDMVGLADW 20

Db 3 tglDMVGL--dW 12

RESULT 15

ID W70361 standard; Protein: 15 AA.

XX W70361;

DT 14-DEC-1998 (first entry)

DE Epitope tag WTGLD, open reading frame 1.

XX Target nucleotide: epitope; gene tagging; epitope tagging.

OS Synthetic.

PN WO9826094-A1.

XX 18-JUN-1998.

PF 09-DEC-1997; 97WO-US22472.

PR 09-DEC-1996; 96US-0762106.

PA (JARV/) JARVIK J W.

XX Jarvik JW;

DR WPI: 1998-348546/30.

XX N-PSDB; V33280.

PT Epitope tagging of genes, transcripts and polypeptides - using an
PT oligonucleotide comprising a nucleotide sequence encoding the
PT epitope independently of the reading frame of the nucleotide
PT sequence

PS Disclosure: Page 15; 46pp; English.

CC The epitope tag is encoded by the nucleotide sequence independently of
CC the reading frame. The epitope and methods disclosed can be used for the
CC tagging of genes, transcripts and polypeptides. They can be used for
CC e.g. discovering new genes, determining the size and abundance of
CC proteins produced by newly discovered genes, tracking the movement of
CC proteins within cell membranes, monitoring receptor binding and
CC internalisation of exogenous proteins, identifying the components of
CC functional protein complexes, purifying proteins, discovering the
CC function of proteins. The products, methods and uses are particularly
CC used for proteins that are unstable, are difficult to purify, or share
CC epitopes with a number of other proteins. The methods overcome the
CC inefficiency of epitope tagging caused by reading frame obstacles and
CC orientation obstacles.

CC Sequence 15 AA;

Query Match 33.0%; Score 36; DB 19; Length 15;
Best Local Similarity 66.7%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

OY 9 TGLDMVGLADW 20
||| |
Db 2 tgldevgl--dw 11

RESULT 16

R65046 R65046 standard; Peptide; 22 AA.

AC R65046;

DT 06-OCT-1995 (first entry)

DE Random biotinylation peptide 15.

XX biotinylation; peptide; recombinant; fusion protein; small;

KW specific; defined; purification; BIRA; enzyme; biotin.

XX Synthetic.

OS Key Location/Qualifiers

FT Modified-site 16

FT /note="biotin-Lys"

PN W09504069-A.

PD 09-FEB-1995.

PF 28-JUL-1994; 94WO-US08528.

PR 30-JUL-1993; 93US-0099991.

PA (AFFY-) AFFYMAX TECHNOLOGIES NV.

PI Schatz PJ;

DR WPI; 1995-090609/12.

PT Production of biotinylated proteins by expression of a
recombinant DNA vector - which encodes a fusion protein
comprising a protein and a biotinylated peptide.

PS Claim 10; Page 136; 146pp; English.

CC A library of small, efficient peptide biotinylation sequences (R65032-46)
CC was generated by using a generic peptide (R65020) and a system known as
CC the "peptides on plasmids" system. At some positions in the sequences,
CC no clear consensus is apparent. At other residues, however, clear trends
CC emerge. A protein can be biotinylated by constructing a recombinant DNA
CC expression vector encoding a fusion protein, comprising a protein and a
CC biotinylation peptide. A host cell, eg. E. coli is transformed with the
CC vector and is cultured in the presence of biotin and a biotinylation
CC enzyme, eg. BIRA.

CC Sequence 22 AA;

Query Match 33.0%; Score 36; DB 16; Length 22;
Best Local Similarity 38.9%; Pred. No. 6.5;
Matches 7; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

OY 7 LSTGLD---MVGLADW 20
: |||| : | : |
Db 2 metgidlrplltgmxdw 19

RESULT 17

W46649

ID W46649 standard; peptide; 22 AA.

AC W46649;

DT 28-MAY-1998 (first entry)

DE Biotinylation peptide isolated from random library 1.

XX Biotinylation peptide; biotinylation enzyme; biotin-protein ligase;

KW BIRA; biotin ligase; biotin; purification; immobilisation; labelling;

KW detection; protein.

XX Synthetic.

OS US5723584-A.

PN 03-MAR-1998.

PD 03-FEB-1995; 95US-0383753.

PR 30-JUL-1993; 93US-0099991.

PA (AFFY-) AFFYMAX TECHNOLOGIES NV.

PI Schatz PJ;

DR WPI; 1998-178542/16.

PT Peptide(s) that can be biotinylated by biotin ligase - and fusion
proteins containing them

PS Claim 3; Column 66; 33pp; English.

CC Peptides W46648-49 are non-naturally occurring biotinylation peptides,
CC derived from a library constructed to express peptides of the generic
CC sequence W46623. The library was constructed using oligonucleotides
CC V6118-20. The peptides contain a biotinylatable sequence motif,
CC recognised by a biotinylation enzyme, e.g. biotin-protein ligase (BIRA).
CC The C or N terminus of the peptides can be covalently coupled to a
CC protein that is incapable of being biotinylated by a biotin ligase.
CC The peptides can be biotinylated in vitro or in vivo, especially with
CC BIRA biotin ligase, and used for the purification, immobilisation,
CC labelling or detection of proteins.

CC Sequence 22 AA;

Query Match 33.0%; Score 36; DB 19; Length 22;
Best Local Similarity 38.9%; Pred. No. 6.5;
Matches 7; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

OY 7 LSTGLD---MVGLADW 20
: |||| : | : |
Db 2 metgidlrplltgmxdw 19

RESULT 18

T29339 T29339 standard; peptide; 22 AA.

AC Y29339;

DT 28-SEP-1999 (first entry)

DE Biotinylation peptide SEQ ID NO:28.

XX Biotinylation enzyme; biotin-protein ligase; BIRA; labelling;

KW biotin carboxyl carrier protein; BCCP; Escherichia coli;

KW fusion protein; identification; purification; diagnosis; research;

XX Escherichia coli.

OS Synthetic.

XX

PN US5932433-A.
 XX 03-AUG-1999.
 PD 28-OCT-1997; 97US-0959512.
 PF 03-FEB-1995; 95US-0383753.
 XX 30-JUL-1993; 93US-0099991.
 PR 28-OCT-1997; 97US-0959512.
 XX
 PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
 XX
 PI Schatz PJ;
 XX
 DR WPI: 1999-457113/38.
 XX
 PT Identification and biotinylation of proteins synthesized by
 PT recombinant DNA techniques in vivo
 XX
 PS Claim 6: Column 66; 35pp; English.
 XX
 CC The present invention describes a method for the identification and
 CC biotinylation of proteins synthesized by recombinant DNA techniques
 CC in vivo with a biotinylation peptide of less than 50 amino acids. The
 CC method comprises: (a) on the surface of a substrate, providing a fusion
 CC protein comprising a recombinant protein and a peptide defined by the
 CC sequence given in Y24492, where the peptide is capable of being
 CC biotinylated by a biotin ligase at the lysine residue adjacent to
 CC position 8 and is 13-50 aa in length; (b) in a predefined region of the
 CC surface of the substrate, contacting the fusion protein with an enzyme;
 CC and (c) determining whether the fusion protein has been biotinylated.
 CC The method allows the identification and purification of biotinylated
 CC enzymes e.g. Birt. The method is also useful in research and diagnostic
 CC applications. The method uses small but specific peptides, allowing the
 CC labelling of a protein at a defined site, this provides improved
 CC immunoblotting and avoids the use of antibodies. Y24493 to Y24548, and
 CC Y29239 to Y29312 represent specifically claimed examples of
 CC biotinylation peptides for use in the method of the invention.
 CC
 XX
 SQ Sequence 22 AA:
 XX
 Query Match 33.0%; Score 36; DB 20; Length 22;
 Best Local Similarity 38.9%; Pred. No. 6.5;
 Matches 7; Conservative 3; Mismatches 4; Indels 4; Gaps 1;
 OY 7 LSTGLD---MVGLAADW 20
 Db 2 metgldlrplltgmkmnw 19
 XX
 RESULT 19
 Y79295
 ID Y79295 standard; Peptide: 15 AA.
 XX
 AC Y79295;
 XX
 DT 18-JUL-2000 (first entry)
 XX
 DE Human prorenin profragment immunogenic antigen.
 XX
 KW Human; prorenin; renin; antibody; immunogen; antigen.
 XX
 OS Homo sapiens.
 XX
 PN EP994187-A1.
 XX
 PD 19-APR-2000.
 XX
 PF 13-OCT-1999; 99EP-0308080.
 XX
 PR 13-OCT-1998; 98JP-0291124.
 XX

PA (TOKU) TOKIMA CHEM IND CO LTD.
 PA (ISHI/) ISHIDA Y.
 XX
 PI Murakami K, Nakamura Y, Suzuki F, Ishida Y;
 XX
 DR WPI: 2000-273528/24.
 XX
 PT Novel renin-active complex comprising human prorenin and at least one
 PT anti-peptide antibody with specific affinity to parts of the human
 PT prorenin profragment -
 XX
 PS Example 2; Page 6; 18pp; English.
 XX
 CC This peptide is a human prorenin profragment immunogenic antigen
 CC comprising amino acid residues 27-41 of the profragment (see also
 CC Y79293). The peptide was produced by solid-phase synthesis, with
 CC addition of a C-terminal Cys residue, and used to raise
 CC anti-peptide antibodies in New Zealand white rabbit. The invention
 CC provides a novel renin-active complex (I) formed from human
 CC prorenin and an anti-peptide antibody that specifically recognizes
 CC either (a) an amino acid sequence of at least 15 residues in the 33
 CC residues between Ile11 and Arg43 of human prorenin profragment, or
 CC (b) an amino acid sequence comprising residues Ile11-Arg26 of human
 CC prorenin profragment, or (c) an amino acid sequence comprising
 CC residues Gly27-Met41 of human prorenin profragment. In all cases,
 CC (I) may also include a second anti-peptide antibody that specifically
 CC recognizes a sequence of 11 amino acids comprising Leu1-Ile11 of
 CC human prorenin profragment. (I) is useful as a renin substrate.
 CC Greatly enhanced renin activity can be obtained by using a mixed
 CC antibody consisting of equimolar amounts of anti-peptide antibodies.
 CC
 XX
 SQ Sequence 15 AA:
 XX
 Query Match 29.4%; Score 32; DB 21; Length 15;
 Best Local Similarity 45.5%; Pred. No. 21;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 10 GLDMVGLAADW 20
 Db 1 gydmvrlgpew 11
 XX
 RESULT 20
 Y67045
 ID Y67045 standard; Peptide: 15 AA.
 XX
 AC Y67045;
 XX
 DT 15-MAR-2000 (first entry)
 XX
 DE B.lentus protease peptide F10.
 XX
 KW Subtilisin; allergy; human; bacterium; protease; epitope; detergent;
 KW cosmetic; textile; pet food industry; debridement treatment.
 XX
 OS Bacillus lentus.
 XX
 PN WO9953078-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 15-APR-1999; 99WO-US08177.
 XX
 PR 15-APR-1998; 98US-0060854.
 XX
 PA (GEWY) GENENCOR INT INC.
 XX
 PI Estell D;
 XX
 DR WPI: 2000-061971/05.
 XX
 PT Modified non-human protease having reduced allergenicity -

XX
PS Example 2; Fig 7; 38pp; English.
XX
CC Peptides Y66983-Y67070 represent peptides derived from the *Bacillus*
CC *lentus* subtilisin protein. The peptides were used for epitope mapping of
CC the bacterial subtilisin. The invention relates to a method of reducing
CC the allergenicity of a non-human protein, especially a bacterial protease
CC such as subtilisin, by identifying an epitope on the non-human protein
CC and replacing it by an analogous region from the human protein. The
CC method is useful for producing proteins, including proteases, that are
CC less likely to cause allergic reactions. Thus the invention can be
CC employed for example in detergents, cosmetics, textile treatment, and
CC pet food industries. The human subtilisin can be used pharmaceutically
CC for debridement treatments.
CC
CC Sequence 15 AA:
SO

Query Match 29.4%; Score 32; DB 21; Length 15;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 FNQSTGLDMV 14
1:1 111:1
Db 3 fsqygaqidv 13

RESULT 21
Y54682
ID Y54682 standard; peptide; 15 AA.
XX
AC Y54682;
XX
DT 04-FEB-2000 (first entry)
XX
DE B. *lentus* protease fragment peptide F10.
XX
KW Subtilisin; protease variant; precursor protease; cleaning composition;
KW detergent; liquid soap application; dish-care formulation; animal feed;
KW contact lens cleaning solution; peptide hydrolysis; waste treatment;
KW cosmetic formulation; fusion-cleavage enzyme; protein production.
XX
OS *Bacillus lentus*.
XX
PN W09953038-A2.
XX
PD 21-OCT-1999.
XX
PF 14-APR-1999; 99WO-US08253.
XX
PR 15-APR-1998; 98US-0060872.
XX
PA (GEMV) GENENCOR INT INC.
XX
PI Estell DA, Harding FA;
XX
DR WPI: 2000-013100/01.
XX
PT New mutant proteins having lower allergenic response in humans useful
PT in cleaning compositions, animal feed and treating textiles -
XX
PS Example 2; Fig 6a; 43pp; English.
XX
CC This sequence is a fragment of the *Bacillus lentus* protease. The
CC invention relates to protease variants comprising a substitution at least
CC 1 of the positions in a precursor protease corresponding to K170D, Y171Q
CC and/or S173D of the *Bacillus amyloliquefaciens* subtilisin. The protease
CC variant can be used in a cleaning composition (e.g. in detergents, in bar
CC or liquid soap application, dish-care formulations, or contact lens
CC cleaning solutions), in animal feed and for treating textiles
CC (e.g. treating wool to prevent felting). The protease variant may also be
CC used for peptide hydrolysis, waste treatment, cosmetic (e.g. skin care)
CC formulations, or as fusion-cleavage enzymes in protein production. The

CC proteins (including enzymes) with reduced antigenicity may be used with
CC significantly less danger of sensitisation for the individuals exposed.
CC Peptides which contain epitopes responsible for initially sensitising an
CC individual may be identified by measuring the proliferation of T-cells
CC due to T-cell epitope recognition.
CC
CC Sequence 15 AA:
SO

Query Match 29.4%; Score 32; DB 21; Length 15;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 FNQSTGLDMV 14
1:1 111:1
Db 3 fsqygaqidv 13

RESULT 22
W13198
ID W13198 standard; Protein; 18 AA.
XX
AC W13198;
XX
DT 19-MAY-1997 (first entry)
XX
DE Fragment of p53 binding protein WBPI.
XX
KW p53; binding protein; WBPI; antibody; diagnosis; treatment;
KW neoplastic; pre-neoplastic; disease; agonist; antagonist;
KW augmentation; inhibition; complex formation; neoplasia; apoptosis;
KW reperfusion; injury; myocardial infarction; stroke; AIDS;
KW traumatic brain; neurodegenerative; aging; ischaemia; toxemia;
KW infection; hepatitis; probe; genetic; forensic identification;
KW fragment.
XX
OS Homo sapiens.
XX
PN W09514777-A1.
XX
PD 01-JUN-1995.
XX
PF 21-NOV-1994; 94WO-US13499.
XX
PR 22-NOV-1993; 93US-0156571.
XX
PA (ONYX-) ONYX PHARM.
XX
PI Bischoff JR, Wu L;
XX
DR WPI: 1995-206934/27.
XX
PT New p53-binding polypeptide(s) WBPI and p53UBC - used to develop
PT prods. for screening assays and for use in diagnosis and therapy of
PT diseases, esp. neoplasia
XX
PS Claim 8; Page 38; 90pp; English.
XX
CC The present sequence is a fragment of the p53 binding protein WBPI,
CC which can be used to generate antibodies for the diagnosis of
CC (pre)neoplastic diseases, and WBPI (ant)agonists. The (ant)agonists
CC augment or inhibit the formation of p53:WBPI complexes and
CC inhibit neoplasia or apoptosis, useful in the treatment of, e.g.
CC reperfusion injury, myocardial infarction, stroke, traumatic brain
CC injury, neurodegenerative diseases, aging, ischaemia, toxemia,
CC infection, AIDS and hepatitis. The WBPI encoding cDNA, which was
CC isolated from a HeLa cell derived cDNA library using the yeast
CC two-hybrid system, can be used to generate probes for the diagnosis
CC of (pre)neoplastic pathological conditions and genetic diseases,
CC and the forensic identification of human individuals.
XX
SO Sequence 18 AA;

Query Match 28.4%; Score 31; DB 16; Length 18;
Best Local Similarity 33.3%; Pred. No. 41;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 PRYFNQSTGLDMVG 15
1 : 1 : 1 : 1 :
Db 1 pgqvgqissmdisg 15

RESULT 23

W57351
ID W57351 standard; peptide; 18 AA.

AC W57351;

DT 11-AUG-1998 (first entry)

DE Human WBPI immunogenic peptide SEQ ID NO:83.

XX Human; WBPI, p53; cancer; interacting protein; screen; diagnosis;
XX genetic disease; forensic identification; nuclear phosphoprotein;
XX cellular proliferation; neoplastic transformation; p53UDC; immunogen.

OS Homo sapiens.

PN US5756669-A.

PD 26-MAY-1998.

PF 02-MAR-1995; 95US-0399696.

PR 02-MAR-1995; 95US-0399696.

PR 22-NOV-1993; 93US-0156571.

PR 21-NOV-1994; 94WO-US13499.

PA (ONYX-) ONYX PHARM INC.

PI Bischoff JR, Wu L;

DR WPI; 1998-321626/28.

PT WBPI, p53-interacting protein - useful screening agents for

PT treatments of p53 related cancers

PS Disclosure; Column 24; 68pp; English.

XX The present sequence represents an immunogenic peptide from WBPI
XX (cellular proliferation and neoplastic transformation) polypeptide,
XX which can be used to screen bacteriophage antibody display libraries or
XX to immunize a rabbit. WBPI is a p53-interacting protein that can be used
XX to screen for agents for diagnosis or treatment of cancer and genetic
XX diseases and for forensic identification of human individuals. The p53
XX protein is a nuclear phosphoprotein involved in control of cellular
XX proliferation, and mutations in it are associated with human cancers.

XX Sequence 18 AA;

Query Match 28.4%; Score 31; DB 19; Length 18;
Best Local Similarity 33.3%; Pred. No. 41;

Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 PRYFNQSTGLDMVG 15
1 : 1 : 1 : 1 :
Db 1 pgqvgqissmdisg 15

RESULT 24

W79452
ID W79452 standard; peptide; 19 AA.

AC W79452;

XX 01-AUG-2000 (first entry)
XX Rat kynurenine aminotransferase variant (aa16-34).
XX

XX Rat; kynurenine aminotransferase; metabolic X syndrome;
XX hypertension; stroke; diabetes; insulin resistance; obesity;
XX dyslipidemia; anorectic; hypotensive; cerebroprotective;
XX vasotropic; antidiabetic; antileptic; gene therapy; diagnosis;
XX mutant; muten; chromosome 3.

OS Rattus sp.

FX Key Location/Qualifiers

FT Peptide 1..19

FT /note- "Corresponds to residues 16-34 of

FT full-length protein"

FT Misc-difference 12

FT /note- "replaces Glu in wild-type"

XX W0200018918-A2.

XX 06-APR-2000.

XX 28-SEP-1999; 99WO-US22494.

XX 28-SEP-1998; 98US-0161939.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA;

XX WPI; 2000-303450/26.

XX Novel genes encoding e.g. CD36, SGLT2, and kynurenine aminotransferase,

XX which are predictive and therapeutic for stroke, hypertension, diabetes

XX and obesity

XX Claim 2; Fig 4B; 79pp; English.

XX This sequence represents amino acids 16-34 of kynurenine
XX aminotransferase (KAT) from a spontaneously hypertensive rat (SHR).
XX This region of KAT carries an amino acid substitution (E27G) when
XX compared to the corresponding region in control Wistar Kyoto (WKY)
XX rats (see Y79453). The mutation changes the charge of a residue
XX that is conserved across all known KAT homologues, from human to
XX Caenorhabditis elegans. The invention discloses genes whose expression
XX is modulated in SHR rat, which is an animal model for human
XX metabolic X syndrome. Heart, brain, fat, liver and kidney tissues
XX from SHR, stroke-prone SHR (SHR-SP) and WKY animals were analysed
XX using GeneCalling methodology to facilitate the identification and
XX characterization of genes which are differentially expressed in the
XX SHR and SHR-SP, as compared with WKY rats. Genes encoding KAT,
XX sodium dependent glucose cotransporter, CD36, aldolase A, atrial
XX natriuretic factor, alpha-cardiac myosin and alpha-tubulin were
XX identified as being potentially associated with hypertension.

XX Obesity and insulin resistance. Rat was increased in abundance in
XX SHR and SHR-SP kidney, and was mapped within an SHR-SP
XX administering kynurenine acid. Hypotension may be treated by
XX of the invention, including isolated proteins comprising the present
XX sequence, are used for treating, preventing and diagnosing ischemic
XX and metabolic diseases and disorders, such as stroke, hypertension,
XX diabetes and obesity, especially insulin resistivity, dyslipidemia
XX and ischemic stroke (all claimed). The polynucleotides may also be
XX used in gene therapy and antisense therapy protocols.

XX Sequence 19 AA;

Query Match 28.4%; Score 31; DB 21; Length 19;
Best Local Similarity 46.2%; Pred. No. 43;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 4 FNQLSTGLDMVGL 16
 | : | : | : | : |
 Db 6 fgkltkygdvnl 18

RESULT 25

Y91936
 ID Y91936 standard; Peptide: 20 AA.

AC Y91936;

DT 19-JUL-2000 (first entry)

DE HsRec2 peptide residues 153-172.

KM HsRec2: serine kinase; homologous pairing; strand transfer; RAD51;
 KW phosphorylation; cyclin E; p53; cell cycle; substrate.

OS Synthetic.

PN WO200017329-A1.

PD 30-MAR-2000.

PF 17-SEP-1999; 99WO-US21642.

PR 21-SEP-1998; 98US-0157603.

PA (UYJE-) UNIV JEFFERSON THOMAS.
 (CORR) CORNELL RES FOUND INC.
 (KIME-) KIMERGEN INC.

PI Havre PA, Rice MC, Holloman WK, Kmlec EB;

DR WPI; 2000-283562/24.

PT Phosphorylating a serine-containing substrate by incubating it with
 PT adenosine triphosphate and Rec2 kinase and measuring the level of
 PT phosphorylation, useful for discovering specific antagonists or
 PT agonists of Rec2

PS Example 3; Page 14; 41pp; English.

CC The present sequence is hsrRec2 peptide, residues 153-172, a substrate
 CC which was not phosphorylated by hsrRec2, a human serine kinase. HsrRec2
 CC is in the same supergene family as the mammalian protein having
 CC homologous pairing and strand transfer activities, RAD51 and was
 CC isolated because of its homology to the homologous pairing and strand
 CC transfer protein of *Ustilago maydis*. In particular, hsrRec2
 CC phosphorylates several proteins that control the cell cycle, especially
 CC cyclin E and p53. The invention permits the phosphorylation of cell
 CC cycle control proteins at sites that are physiologically relevant. The
 CC invention can be practiced with either murine or human Rec2 or a mutain
 CC or chimera of these proteins. In particular the mutain has the sequence
 CC of a Rec2 kinase containing other than a Tyr at position 163. The
 CC invention comprises a method of phosphorylating a serine-containing
 CC substrate comprising incubating the substrate with ATP (adenosine
 CC triphosphate) and Rec2 kinase or a mammalian Rec2 and measuring the level
 CC of phosphorylation. The method is useful for discovering compounds which
 CC are specific antagonists or agonists of Rec2.

SO Sequence 20 AA;

Query Match 28.4%; Score 31; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 46;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRYFN 5

Db 9 prYfn 13

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:55:27 ; Search time 93.77 Seconds
(without alignments)
24.999 Million cell updates/sec

Title: US-08-981-824-6

Sequence: 1 TVEIAPVLEVTLLKMR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PODONT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29.3	19	2	Q9S625	Q9S625 prochloroto
2	29.3	23	2	Q9S630	Q9S630 prochloroto
3	27.3	18	8	Q9ZYW9	Q9ZYW9 agathielia
4	27.3	23	11	Q63334	Q63334 rattus norv
5	26.3	19	2	Q9X3F0	Q9X3F0 prochloroto
6	26.3	25	11	Q63992	Q63992 rattus norv
7	25.3	18	6	Q9ZTOR	Q9ZTOR mamuthus p
8	25.3	18	8	Q9ZTV8	Q9ZTV8 aspiiota sp
9	25.3	19	8	Q9ZYV7	Q9ZYV7 myrmecia fo
10	25.3	22	12	Q34195	Q34195 ethrichia r
11	25.3	15	1	Q9QPE6	Q9QPE6 hepatitis c
12	24.2	15	10	Q9UWH4	Q9UWH4 pyrococcus
13	24.2	18	7	Q9S8K7	Q9S8K7 copitis japo
14	24.2	18	12	Q9TNP1	Q9TNP1 mus sp. mhc
15	24.2	22	8	Q41588	Q41588 human immun
16	24.2	22	8	Q99389	Q99389 ovis aries
17	24.2	23	4	Q9NZ67	Q9NZ67 homo sapien
18	24.2	24	4	Q13660	Q13660 homo sapien
19	24.2	25	2	Q9RLN0	Q9RLN0 neisseria m

20	24.2	25	8	Q9MQK4	Q9MQK4 rupicapra r
21	23.5	24	6	Q28200	Q28200 bos taurus
22	23.2	18	2	Q50076	Q50076 mycobacteri
23	23.2	18	8	Q9ZYX0	Q9ZYX0 blacus sp.
24	23.2	25	11	Q9JTK04	Q9JTK04 mus musculu
25	22.5	15	2	Q9RK02	Q9RK02 salmoneilla
26	22.2	12	8	Q9XNR6	Q9XNR6 pylatella 1
27	22.2	18	12	Q87589	Q87589 chimpanzee
28	22.2	18	12	Q87591	Q87591 chimpanzee
29	22.2	18	12	Q87593	Q87593 chimpanzee
30	22.2	19	5	Q20438	Q20438 caenorhabdi
31	22.2	19	8	Q9ZYX2	Q9ZYX2 centistes s
32	22.2	19	8	Q9ZYX7	Q9ZYX7 microplitis
33	22.2	19	10	Q9S8E9	Q9S8E9 zea mays (m
34	22.2	20	11	Q9R5D0	Q9R5D0 pseudomonas
35	22.2	20	13	Q9OV83	Q9OV83 rattus norv
36	22.2	20	13	Q9RP2	Q9RP2 gallus gall
37	22.2	22	5	P82108	P82108 mythima un
38	22.2	22	7	Q9MK47	Q9MK47 oryzias lat
39	22.2	22	13	Q9PS00	Q9PS00 microgogon
40	22.2	23	4	Q16137	Q16137 homo sapien
41	22.2	24	8	Q9MQK1	Q9MQK1 capra ibex
42	22.2	25	5	Q94680	Q94680 polyandroca
43	22.2	25	8	Q9ZYX1	Q9ZYX1 dloepilius s
44	21.7	25	4	Q16092	Q16092 homo sapien
45	21.2	11	2	Q9R5R4	Q9R5R4 mycoplasma

ALIGNMENTS

RESULT 1
ID Q9S625 PRELIMINARY; PRT; 19 AA.
AC Q9S625;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CYTOCHROME B (FRAGMENT).
GN PETB.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID-1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630(1998).
DR EMBL; AF070177; AAD20793.1; -
FT NON_TER 1
SQ SEQUENCE 19 AA; 2207 MW; 63744292AB2D51B4 CRC64;

Query Match 29.3%; Score 29; DB 2; Length 19;
Best Local Similarity 33.3%; Pred. No. 3.6e+02;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 7 VEVLEVTLLK 18
DB 1 VEVLEVTLLK 12
RESULT 2
ID Q9S630 PRELIMINARY; PRT; 23 AA.
AC Q9S630;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CYTOCHROME B (FRAGMENT).
GN PETB.

OS Prochlorococcus sp.
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
 OC Prochlorococcus.
 CX NCBI_TaxID=1220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Urdach E., Chisholm S.W.;
 RT "Genetic diversity in Prochlorococcus populations flow cytometrically
 sorted from the Sargasso Sea and Gulf Stream."
 RL Limnol. Oceanogr. 43:1615-1630(1998).
 DR EMBL: AF070199; AAD23242.1;
 FT NON_TER 1 1
 SQ SEQUENCE 23 AA; 2720 MW; 1474425C2B9CEE6 CRC64;

Query Match 29.3%; Score 27; DB 2; Length 23;
 Best Local Similarity 33.3%; Pred. No. 4.3e+02;
 Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 7 VFVLEVTYTLKK 18
 DB 5 VFVLMHFLMIRK 16

RESULT 3
 OQZYW9 PRELIMINARY; PRT; 18 AA.
 AC OQZYW9;
 DT 01-NOV-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE CYTOCHROME OXIDASE II (FRAGMENT).
 OS Agathella sp.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 OC Ichneumonidae; Braconidae; Agathidinae; Agathella.
 CX NCBI_TaxID=64848;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99152621; PubMed=10028295;
 RA Dowson M., Austin A.D.;
 RT "Evolutionary dynamics of a mitochondrial rearrangement 'hot spot' in
 the Hymenoptera."
 RL Mol. Biol. Evol. 16:298-309(1999).
 DR EMBL: AF034592; AAC79740.1;
 KW Mitochondrion.
 FT NON_TER 1 1
 SQ SEQUENCE 18 AA; 2238 MW; FAB24589081FCCD6 CRC64;

Query Match 27.3%; Score 27; DB 8; Length 18;
 Best Local Similarity 50.0%; Pred. No. 7.3e+02;
 Matches 5; Conservative 3; Mismatches -2; Indels 0; Gaps 0;

OY 9 VVLEVTYTLKK 18
 DB 1 IVLESTELTK 10

RESULT 4
 O63334 PRELIMINARY; PRT; 23 AA.
 AC O63334;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE ALPRA-2-MACROGLOBULIN (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-WISTAR;
 RX MEDLINE=85207604; PubMed=2581948;
 RA Northmann W., Heisig M., Kunz D., Heinrich P.C.;
 RT "Molecular cloning of cDNA sequences for rat alpha 2-macroglobulin and
 measurement of its transcription during experimental inflammation."
 RL J. Biol. Chem. 260:6200-6205(1985).
 DR EMBL: M84369; AAA41594.1;
 DR INTERPRO: IPR001599;
 DR PFM: PF00207; A2M; 1.
 FT NON_TER 1 1
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2526 MW; E2D721FB1B223876 CRC64;

Query Match 27.3%; Score 27; DB 11; Length 23;
 Best Local Similarity 62.5%; Pred. No. 9.4e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 8 FVLEVTY 15
 DB 8 YVLAAYLT 15

RESULT 5
 OQX3F0 PRELIMINARY; PRT; 19 AA.
 AC OQX3F0;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE CYTOCHROME B (FRAGMENT).
 GN PETB.
 OS Prochlorococcus sp.
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
 OC Prochlorococcus.
 CX NCBI_TaxID=1220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Urdach E., Chisholm S.W.;
 RT "Genetic diversity in Prochlorococcus populations flow cytometrically
 sorted from the Sargasso Sea and Gulf Stream."
 RL Limnol. Oceanogr. 43:1615-1630(1998).
 DR EMBL: AF070148; AAD20767.1;
 FT NON_TER 1 1
 SQ SEQUENCE 19 AA; 2247 MW; 4E744292BAD51A3 CRC64;

Query Match 26.3%; Score 26; DB 2; Length 19;
 Best Local Similarity 25.0%; Pred. No. 1.1e+03;
 Matches 3; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 7 VFVLEVTYTLKK 18
 DB 1 VFVLMHFLMIRK 12

RESULT 6
 O63992 PRELIMINARY; PRT; 25 AA.
 AC O63992;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE HOX3.2 HOMEOBOX HOMOLOG (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94271262; PubMed=7911662;
 RA Iimura T., Oida S., Takeda K., Maruoka Y., Sasaki S.;

RT "Changes in homeobox-containing gene expression during ectopic bone formation induced by bone morphogenetic protein."
 RL Biochem. Biophys. Res. Commun. 201:980-987(1994).
 DR EMBL: S71317; AAB31008.1; -
 DR HSSP: P02833; 9ANT.
 DR INTERPRO: IPR001356; -
 DR PIRAM: PF00046; homeobox; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT NON_TER 1
 SQ SEQUENCE 25 AA; 3157 MW; 42C4F3D514AD2304 CRC64;

Query Match 26.3%; Score 26; DB 11; Length 25;
 Best Local Similarity 54.5%; Pred. No. 1.5e+03;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 YEAPVFLLE 12
 DB 12 YEVARVNLTE 22

RESULT 7
 ID 09TOR9 PRELIMINARY; PRT; 18 AA.
 AC 09TOR9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DE 01-MAY-2000 (TREMBlrel. 13, last annotation update)
 DE VON WILLEBRAND FACTOR (FRAGMENT).
 OS Mammuthus primigenius (Siberian woolly mammoth).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Proboscidea; Elephantidae; Mammuthus.
 OX NCBI_TaxID=37349;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20022977; PubMed=10555277;
 RA Greenwood A.D.; Capelli C.; Possner G.; Paabo S.;
 RT "Nuclear DNA sequences from late Pleistocene megafauna."
 RL Mol. Biol. Evol. 16:1466-1473(1999).
 DR EMBL: AF154874; AAF12750.1; -
 DR EMBL: AF154873; AAF12749.1; -
 FT NON_TER 1
 FT NON_TER 18
 SQ SEQUENCE 18 AA; 1914 MW; DFCB484B41F69236 CRC64;

Query Match 25.3%; Score 25; DB 6; Length 18;
 Best Local Similarity 60.0%; Pred. No. 1.6e+03;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 5 APVFLLEV 14
 DB 1 APVSVTPV 10

RESULT 8
 ID 09ZYV8 PRELIMINARY; PRT; 18 AA.
 AC 09ZYV8;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, last annotation update)
 DE CYTOCHROME OXIDASE II (FRAGMENT).
 OS Aspilota sp.
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 OC Ichneumonidae; Braconidae; Alysiinae; Aspilota.
 OX NCBI_TaxID=61200;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99152621; PubMed=10028295;

RA Dowton M., Austin A.D.;
 RT "Evolutionary dynamics of a mitochondrial rearrangement 'hot spot' in the Hymenoptera."
 RL Mol. Biol. Evol. 16:298-309(1999).
 DR EMBL: AF034603; AAC79751.1; -
 KW Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 18 AA; 2181 MW; F83846FDAE8BDCD6 CRC64;

Query Match 25.3%; Score 25; DB 8; Length 18;
 Best Local Similarity 55.6%; Pred. No. 1.6e+03;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 9 VLEEVYTLK 17
 DB 1 IVLESVSLK 9

RESULT 9
 ID 09ZYV7 PRELIMINARY; PRT; 19 AA.
 AC 09ZYV7;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
 DE 01-MAY-1999 (TREMBlrel. 10, last annotation update)
 DE CYTOCHROME OXIDASE II (FRAGMENT).
 OS Myrmica forficata.
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Aculeata;
 OC Formicidae; Formicidae; Myrmecia.
 OX NCBI_TaxID=32387;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99152621; PubMed=10028295;
 RA Dowton M., Austin A.D.;
 RT "Evolutionary dynamics of a mitochondrial rearrangement 'hot spot' in the Hymenoptera."
 RL Mol. Biol. Evol. 16:298-309(1999).
 DR EMBL: AF034604; AAC79752.1; -
 KW Mitochondrion.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 19 AA; 2370 MW; 25AE99B2B1F53E6B CRC64;

Query Match 25.3%; Score 25; DB 8; Length 19;
 Best Local Similarity 33.3%; Pred. No. 1.7e+03;
 Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 9 VLEEVYTLK 17
 DB 1 IVLEVTNIK 9

RESULT 10
 ID 034195 PRELIMINARY; PRT; 22 AA.
 AC 034195;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, last annotation update)
 DE GROES. (FRAGMENT).
 GN GROES.
 OS Ehrlichia risticii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichiae; Ehrlichia.
 OX NCBI_TaxID=950;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97373904; PubMed=9230387;
 RA Sumner J.W., Nicholson W.L., Massung R.F.;
 RT "PCR amplification and comparison of nucleotide sequences from the

RT groESL heat shock operon of *Escherichia* species."
 RL J. Clin. Microbiol. 35:2087-2092(1997).
 DR EMBL: U96732; AAB5632.1; -
 DR INTERPRO: IPR001476; -
 DR Pfam; PF00166; cpn10; 1.
 FT NON_TER
 SQ SEQUENCE 22 AA: 2504 MW: 210959075191620 CRC64;

Query Match 25.3%; Score 25; DB 2; Length 22;
 Best Local Similarity 50.0%; Pred. No. 1.9e+03;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 11 LEVYTLKK 18
 :||| :||
 DB 5 VEYVWKE 12

RESULT 11
 O90P62 PRELIMINARY; PRT; 25 AA.
 AC O90P62;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE TRUNCATED NS2.
 OS Hepatitis C virus.
 CC Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 CC Hepacivirus.
 CC NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Leng Y., Chen X.Z., Du Y., Wang H.T.;
 RT "Cloning of NS2 cDNA of Hepatitis C virus (HCV) and complexity of NS2
 RT quasispecies in a HCV carrier."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF175472; AAD50790.1; -
 SQ SEQUENCE 25 AA: 2637 MW: 5684D08FA4DAC202 CRC64;

Query Match 25.3%; Score 25; DB 12; Length 25;
 Best Local Similarity 60.0%; Pred. No. 2.2e+03;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 7 VFVLEVYTL 16
 ||| :||
 DB 13 VFVGLVFTL 22

RESULT 12
 O9UWH4 PRELIMINARY; PRT; 15 AA.
 AC O9UWH4;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE KETO-VALINE-FERREDOXIN OXIDOREDUCTASE ALPHA-SUBUNIT, VOR ALPHA
 DE (FRAGMENT).
 OS Pyrococcus.
 CC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae.
 CC NCBI_TaxID=2260;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-96146528; PubMed-8550513;
 RA Heider J., Mai X., Adams M.W.;
 RT "Characterization of 2-ketoglutarate ferredoxin oxidoreductase, a
 RT new and reversible coenzyme A-dependent enzyme involved in peptide
 RT fermentation by hyperthermophilic archaea."
 RL J. Bacteriol. 178:780-787(1996).
 SQ SEQUENCE 15 AA: 1727 MW: 42363C008C9A9CFA CRC64;

Query Match 24.2%; Score 24; DB 1; Length 15;

Best Local Similarity 44.4%; Pred. No. 1.9e+03;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 11 LEVYTLKK 19
 :||| :||
 DB 1 MEYPIKRV 9

RESULT 13
 O9S8K7 PRELIMINARY; PRT; 15 AA.
 AC O9S8K7;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE S-ADENOSYL-L-METHIONINE:NORCOCLAURINE 6-O-METHYLTRANSFERASE,
 DE NORCOCLAURINE 6-O-METHYLTRANSFERASE (EC 2.1.1.-) (FRAGMENT).
 OS Coplis japonica (Japanese goldthread).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; Ranunculales; Ranunculaceae; Coplis.
 CC NCBI_TaxID=3442;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-95010097; PubMed-7925429;
 RA Sato F., Tsujita T., Katagiri Y., Yoshida S., Yamada Y.;
 RT "Purification and characterization of S-adenosyl-L-methionine:
 RT norcochlorine 6-O-methyltransferase from cultured Coplis japonica
 RT cells."
 RL Eur. J. Biochem. 225:125-131(1994).
 SQ SEQUENCE 15 AA: 1595 MW: 8762BC05939B6C4C CRC64;

Query Match 24.2%; Score 24; DB 10; Length 15;
 Best Local Similarity 40.0%; Pred. No. 1.9e+03;
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 YEIAVFTL 11
 :||| :||
 DB 1 YALXPATL 10

RESULT 14
 O9TNP1 PRELIMINARY; PRT; 18 AA.
 AC O9TNP1;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE MHC CLASS II H 2EK NATURAL LIGAND (FRAGMENT).
 OS Mus sp.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10095;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-96360157; PubMed-8746565;
 RA Schild H., Grunberg U., Pougialis G., Wallny H.J., Kellholz W.,
 RA Stevanovic S., Rammensee H.G.;
 RT "Natural ligand motifs of H-2E molecules are allele specific and
 RT illustrate homology to HLA-DR molecules."
 RL Int. Immunol. 7:1957-1965(1995).
 KV MHC.
 SQ SEQUENCE 18 AA: 1997 MW: A0FD2CD387023F46 CRC64;

Query Match 24.2%; Score 24; DB 7; Length 18;
 Best Local Similarity 71.4%; Pred. No. 2.3e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 12 EYVTLK 18
 :||| :||
 DB 3 EFGTLK 9

RESULT 15
 O41588
 ID 041588 PRELIMINARY; PRT; 18 AA.
 AC 041588;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PC13;
 RX MEDLINE-98105804; PubMed-9445059;
 RA Connor R.I., Korber B.T.M., Graham B.S., Hahn B.H., Ho D.D.,
 Walker B.D., Neumann A.U., Vermund S.H., Westecky J., Jackson S.,
 Fenimore E., Cao Y., Gao F., Kalam S., Kunstman K.J., McDonald D.,
 McWilliams N., Tirkola A., Moore J.P., Wolinsky S.M.;
 RT Immunological and virological analyses of persons infected by human
 immunodeficiency virus type 1 while participating in trials of
 recombinant gp120 subunit vaccines.";
 RL J. Virol. 72:1552-1576(1998).
 DR EMBL: U84837; AAC5876.1; -.
 DR INTERPRO: IPR000777; -.
 DR PFAM: PF00516; GP120; 1.
 DR Envelope protein.
 FT NON_TER 1
 SQ SEQUENCE 18 AA; 2050 MW; 656402F80913D46E CRC64;

Query Match 24.2%; Score 24; DB 12; Length 18;
 Best Local Similarity 36.4%; Pred. No. 2.3e+03;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 3 EIAPEVLEXY 13
 : : : : :
 DB 5 KLRPLCVTLKY 15

RESULT 16
 O99389
 ID 099389 PRELIMINARY; PRT; 22 AA.
 AC 099389;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE ATPASE SUBUNIT 6 (FRAGMENT).
 OS Ovis aries (Sheep).
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PERIPHERAL BLOOD;
 RX MEDLINE-98403943; PubMed-9734874;
 RA Tartaglia M., Saulle E.;
 RT "Rapid communication: nucleotide sequence of porcine and ovine
 rRNA(Lys) and ATPase8 mitochondrial genes.";
 RL J. Anim. Sci. 76:2207-2208(1998).
 DR EMBL: AF039171; AAD05068.1; -.
 DR Mitochondrion.
 FT NON_TER 22
 SQ SEQUENCE 22 AA; 2453 MW; 215F87BDF0E8A1C8 CRC64;

Query Match 24.2%; Score 24; DB 8; Length 22;
 Best Local Similarity 46.2%; Pred. No. 2.8e+03;
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 IAPFVLEXYTL 16
 : : : : :
 DB 10 ITPMGLPLVTL 22
 RESULT 17
 O99267
 ID 099267 PRELIMINARY; PRT; 23 AA.
 AC 099267;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE ATP7B (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wu Z.Y., Wang N., Murong S.X.;
 RT "Missense mutations of exons 14 and 18 of Wilson's disease gene in
 Chinese patients.";
 RL Chung-Hua I Hsueh I Chuan Hsueh Tsa Chih 16:91-93(1999).
 DR EMBL: AF218354; AAF25685.1; -.
 FT NON_TER 1
 FT VARIANT 9
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2483 MW; 23CF54F1868B5DOB CRC64;

Query Match 24.2%; Score 24; DB 4; Length 23;
 Best Local Similarity 38.5%; Pred. No. 3e+03;
 Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 4 IAPFVLEXYTL 16
 : : : : :
 DB 7 VMPVLLGDVATL 19

RESULT 18
 O13660
 ID 013660 PRELIMINARY; PRT; 24 AA.
 AC 013660;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 08, Last annotation update)
 DE MANNOSE 6-PHOSPHATE/INSULIN-LIKE GROWTH FACTOR-II RECEPTOR
 (FRAGMENT).
 OS Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96083596; PubMed-7493029;
 RA De Souza A.T., Hankins G.R., Washington M.K., Otton T.C., Jittle R.L.;
 RT "M6P/IGF2R gene is mutated in human hepatocellular carcinomas with
 loss of heterozygosity.";
 RL Nat. Genet. 11:447-449(1995).
 DR EMBL: S80783; AAB35664.1; -.
 FT NON_TER 1
 SQ SEQUENCE 24 AA; 2799 MW; E6F28C7F43E17602 CRC64;

Query Match 24.2%; Score 24; DB 4; Length 24;
 Best Local Similarity 50.0%; Pred. No. 3.1e+03;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 7 VFVLEXY 14
 : : : : :
 DB 6 VIVLVKY 13

RESULT 19

09RLNO PRELIMINARY: PRT: 25 AA.
 ID 09RLNO
 AC 09RLNO:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE POR. PROTEIN (FRAGMENT).
 GN POR.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 NCBI_TaxID=487;
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=36/93;
 RA Wedge E., Caugant D.A., Musacchio A., Saunders N.B., Zollinger W.D.,
 "Redesignation of a purported P1.15 subtype-specific meningococcal
 RT monoclonal antibody as a P1.19-specific reagent."
 RL Clin. Diagn. Lab. Immunol. 6:639-642(1999).
 DR EMBL; AJ012726; CAA10149.1; -
 RL NON-TER 1 1
 FT NON-TER 25 25
 FT SEQUENCE 25 AA: 2480 MW: 3AD7F47D1C4D8378 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 25;
 Best Local Similarity 35.7%; Pred. No. 3.2e+03;
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 6 PVFVLELYVTLLK 19
 DB 7 PAYVAENGVAR 20

RESULT 20

09MOK4 PRELIMINARY: PRT: 25 AA.
 ID 09MOK4
 AC 09MOK4:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE ATPASE SUBUNIT 6 (FRAGMENT).
 OS Rhipicarpa rupicapra.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Rupicapra.
 OX NCBI_TaxID=34869;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PERIPHERAL BLOOD;
 RA MEDLINE=20104876; PubMed=10641890;
 RA Saule E., Di Pasquale S., Tartaglia M.,
 "Rapid communication: nucleotide sequence of chamois, alpine ibex, and
 RT red deer (Capra) and Arpase8 mitochondrial genes."
 RL J. Anim. Sci. 77:3398-3399(1999).
 DR EMBL; AF104681; AAF43479.1; -
 RN Mitochondrion.
 KW NON-TER 25 25
 FT NON-TER 25 25
 FT SEQUENCE 25 AA: 2778 MW: 6FFDC915FF87BDF CRC64;

Query Match 24.2%; Score 24; DB 8; Length 25;
 Best Local Similarity 46.2%; Pred. No. 3.2e+03;
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 IAPFVLELYVTLL 16
 DB 10 ITPMMGLPLVTL 22

RESULT 21

Q28200 PRELIMINARY: PRT: 24 AA.
 ID Q28200
 AC Q28200:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 08, Last annotation update)
 DE TYROSINASE-RELATED PROTEIN-1 (FRAGMENT).
 GN TRP-1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96314791; PubMed=8759137;
 RA Noneman D., Shubuya H., Johnson G.S.;
 "A BstUI PCR/RFLP in the bovine tyrosinase-related protein-1 (TYRPL)
 RT gene."
 RL Anim. Genet. 27:218-219(1996).
 DR EMBL; L43123; AAB38537.1; -
 RL NON-TER 1 1
 FT NON-TER 24 24
 FT SEQUENCE 24 AA: 2719 MW: C72B279D86390579 CRC64;

Query Match 23.7%; Score 23.5; DB 6; Length 24;
 Best Local Similarity 42.1%; Pred. No. 3.8e+03;
 Matches 8; Conservative 3; Mismatches 5; Indels 3; Gaps 1;

OY 1 TYET---APFVLELYVTLL 16
 DB 5 TYEIQWPSRSFSISGV 23

RESULT 22

050076 PRELIMINARY: PRT: 18 AA.
 ID 050076
 AC 050076:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)
 DE u1400.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1169;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R.;
 RN Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Robison K.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U15183; AAA63011.1; -
 SO SEQUENCE 18 AA: 2183 MW: 951B7CD608037D00 CRC64;

Query Match 23.2%; Score 23; DB 2; Length 18;
 Best Local Similarity 50.0%; Pred. No. 3.4e+03;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 7 VFVLELYVTLL 16
 DB 2 VFELYEHTSL 11

RESULT 23

09ZYX0 PRELIMINARY: PRT: 18 AA.
 ID 09ZYX0
 AC 09ZYX0:
 DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE CYTOCHROME OXIDASE II (FRAGMENT).
 OS Blacus sp.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 OC Ichneumonidae; Braconidae; Blacinae; Blacus.
 OX NCBI_TaxID=64854;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9152621; PubMed=10028295;
 RA Dowton M., Austin A.D.;
 RT "Evolutionary dynamics of a mitochondrial rearrangement 'hot spot' in the Hymenoptera."
 RT Mol. Biol. Evol. 16:298-309(1999).
 RL EMBL; AF034591; AAC79739.1; -.
 DR EMBL; AF034591; AAC79739.1; -.
 KW Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 18 AA; 2269 MW; E826148E78157286 CRC64;

Query Match 23.2%; Score 23; DB 8; Length 18;
 Best Local Similarity 40.0%; Pred. No. 3.4e+03;
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 9 VLEVTLKK 18
 :|:|:|
 Db 1 IVLEVIELNK 10

RESULT 24
 O9JRK04 PRELIMINARY; PRT; 25 AA.
 ID O9JK04
 AC O9JK04
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE FRUCTOSE-1,6-BISPHOSPHATASE (EC 3.1.3.11) (FRAGMENT).
 GN FBPA2.2
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tillmann H., Stein S., Liehr T., Eschrich K.;
 RT "Structure and chromosomal localization of the human and mouse muscle fructose-1,6-bisphosphatase genes."
 RT Gene 247:241-253(2000).
 RL EMBL; AJ243025; CAB90672.1; -.
 DR EMBL; AJ243025; CAB90672.1; -.
 KM Hydrolyase.
 FT NON_TER 1
 FT NON_TER 25
 SQ SEQUENCE 25 AA; 2898 MW; E4A7A6F50F3A31D3 CRC64;

Query Match 23.2%; Score 23; DB 11; Length 25;
 Best Local Similarity 71.4%; Pred. No. 4.7e+03;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 12 EYVTLKK 18
 :|:|:|
 Db 16 EYVOKKK 22

RESULT 25
 O9RQ22 PRELIMINARY; PRT; 15 AA.
 ID O9RQ22
 AC O9RQ22
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)

DE FERRIC HYDROXAMATE UPTAKE PROTEIN (FRAGMENT).
 GN FHUB.
 OS Salmonella typhl.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ISP1820;
 RA Morrow B.J., Graham J.E., Curtiss R. III;
 RT "genomic subtractive hybridization and selective capture of transcribed sequences identify a novel Salmonella typhimurium fimbrial operon and putative transcriptional regulator that are absent from the RT Salmonella typhl genome."
 RT Infect. Immun. 67:5106-5116(1999).
 RL EMBL; AF134977; AAD54116.1; -.
 DR EMBL; AF134977; AAD54116.1; -.
 FT NON_TER 1
 SQ SEQUENCE 15 AA; 1825 MW; 036E36EB6455E616 CRC64;

Query Match 22.7%; Score 22.5; DB 2; Length 15;
 Best Local Similarity 37.5%; Pred. No. 3.4e+03;
 Matches 6; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

OY 5 APVFVLEVTLKKMR 20
 :|:|:|:|
 Db 3 APVFI---YLLRKQSR 15

Search completed: February 5, 2001, 10:55:28
 Job time: 917 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:56:02 ; Search time 32.57 Seconds
(without alignments)
19.831 Million cell updates/sec

Title: US-08-981-824-6

Sequence: 1 TTEIAPVFLVLEVTLMKMR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 1436

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	22.2	6	1	CIP2_MYTED
2	21	21.2	6	1	YCXB_OPOSI
3	20.5	20.7	22	1	MOT1_CANFA
4	20	20.2	15	1	UC20_MAIZE
5	20	20.2	19	1	MIFR_TRISP
6	20	20.2	20	1	AMP_FUSNU
7	20	20.2	23	1	MDH_THEAL
8	19.5	19.7	16	1	PSA1_SPTOL
9	19.5	19.7	25	1	PSA2_SPTOL
10	19	19.2	6	1	CIP1_MYTED
11	19	19.2	6	1	CIP1_MYTED
12	19	19.2	14	1	ULM4_HUMAN
13	19	19.2	14	1	UNO4_CLOPA
14	19	19.2	20	1	MIF_PIG
15	19	19.2	20	1	SUCB_CANFA
16	19	19.2	22	1	NUO7_SOLTU
17	19	19.2	22	1	VGL6_RABVA
18	18.5	18.7	13	1	ATP1_SOLTU
19	18.5	18.7	19	1	PSA1_PEA
20	18	18.2	11	1	CH10_CLOPA
21	18	18.2	15	1	MIFR_BRAJA
22	18	18.2	15	1	MIFR_ONCKE
23	18	18.2	15	1	UC17_MAIZE
24	18	18.2	18	1	AL12_CYPDO
25	18	18.2	20	1	AROC_AMEYE
26	18	18.2	20	1	CPX2_RHORI
27	18	18.2	21	1	GTS2_ASCSU
28	18	18.2	24	1	DCM5_PSECA
29	18	18.2	25	1	CR17_LITXA
30	18	18.2	25	1	COX6_NEDCR
31	18	18.2	25	1	FRHB_METBA
32	17.5	17.7	21	1	SMRP_RAT
33	17.5	17.7	22	1	CH1C_PEA
					IAD1_ENTFA

34	17.5	17.7	24	1	THD2_SALTY
35	17	17.2	10	1	TMOF_AEDAE
36	17	17.2	12	1	TM2A_METMA
37	17	17.2	13	1	ODPA_CANFA
38	17	17.2	14	1	LPW_ECOLI
39	17	17.2	15	1	ACBA_ACTCA
40	17	17.2	15	1	KLOM_LUMTE
41	17	17.2	15	1	UC14_MAIZE
42	17	17.2	15	1	UC19_MAIZE
43	17	17.2	16	1	PGTL_PELAC
44	17	17.2	16	1	UPAB_HUMAN
45	17	17.2	17	1	AP1D_BOMPA

ALIGNMENTS

RESULT 1
CIP2_MYTED
ID CIP2_MYTED STANDARD; PRT; 6 AA.
AC P13737;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CONTRACTION-INHIBITING PEPTIDE II (MIP II).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
RN [1]
RP SEQUENCE.
RC TISSUE-PEDAL GANGLION;
RX MEDLINE-88240357; PubMed-3377776;
RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
RT "Structures and actions of Mytilus inhibitory peptides.";
RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988)
CC -1- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN
CC MUSCLES.
CC -1- SIMILARITY: TO MIP I.
CC PIR: B27696; B27696.
DR PIR: B27696; B27696.
KW Hormone; Amidation.
FT MOD RES
SQ SEQUENCE 6 AA: 621 MW: 72C9C6876BD81000 CRC64;
AMIDATION.
Query Match 22.2%; Score 22; DB 1; Length 6;
Best local Similarity 80.0%; Pred. No. 8.9e+04;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 5 APYFV 9
Db 2 APMEV 6
RESULT 2
YCXB_OPOSI
ID YCXB_OPOSI STANDARD; PRT; 25 AA.
AC P49834;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOHETICAL 3.1 KDA PROTEIN IN PSBJ-TRNE INTERGENIC REGION (ORF25).
OS Odontella sinensis.
OC Chloroplast.
OC Eukaryota; streptophytes; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscaceae; Odontella.
RN [1]
RP SEQUENCE FROM N.A.
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freter U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
Odontella sinensis.";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: 267753; CAA91716.1; -
 KM Chloroplast; Hypothetical protein.
 SQ SEQUENCE 25 AA; 3083 MW; 9F50BB79484E531 CRC64;

Query Match 21.2%; Score 21; DB 1; Length 25;
 Best Local Similarity 36.4%; Pred. No. 1.8e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 7 VFVLEVTYTK 17
 DB 1 MFEISFYIALK 11

RESULT 3
 MOTI_CANFA STANDARD; PRT; 22 AA.
 AC P19863:
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE MOTILIN.
 GN MLN.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=INTESTINE;
 RX MEDLINE=83195948; PubMed=6844663;
 RA Poltras P., Reeve J.R., Jr., Hunziker M.W., Hood L.E., Walsh J.H.,
 RT "Purification and characterization of canine intestinal motilin.";
 RL Regul. Pept. 5:197-208(1983).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE REGULATION OF
 CC INTERDIGESTIVE GASTROINTESTINAL MOTILITY AND INDIRECTLY CAUSES
 CC RHYTHMIC CONTRACTION OF DUODENAL AND COLONIC SMOOTH MUSCLE.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE MOTILIN FAMILY.
 CC PIR: S00189; S00189.
 DR PIR: A60313; A60313.
 KM Hormone.
 FT UNSURE
 SQ SEQUENCE 22 AA; 2685 MW; 4BECB840ABE0639F CRC64;

Query Match 20.7%; Score 20.5; DB 1; Length 22;
 Best Local Similarity 40.0%; Pred. No. 1.9e+03;
 Matches 6; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

OY 6 PVFVLEVTYTKMR 20
 DB 3 PIFTTSE--LQKIR 14

RESULT 4
 UC20_MAIZE STANDARD; PRT; 15 AA.
 AC P80626:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 445)
 DE (FRAGMENT).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.

RN [1]
 RP SEQUENCE.
 RC TISSUE=COLEOPTILE;
 RA Touret P., Roccardi F., Morin C., Damerval C., Huet J.-C.,
 RA Pernollet J.-C., Zivy M., de Vienne D.;
 RT "The maize two dimensional gel protein database: towards an integrated
 RT genome analysis program";
 RL Theor. Appl. Genet. 93:997-1005(1996).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED P1 OF THIS UNKNOWN
 CC PROTEIN IS: 6.1, ITS MW IS: 45.3 KDA.
 CC -1- SIMILARITY: TO ADOMET SYNTHETASES.
 DR MAIZE-2DPAGE:P80626; COLEOPTILE.
 DR MAIZE2B: 123952; -;
 FT NON_TER
 FT NON_TER 15
 SQ SEQUENCE 15 AA; 1597 MW; D2EECE163EB10156 CRC64;

Query Match 20.2%; Score 20; DB 1; Length 15;
 Best Local Similarity 66.7%; Pred. No. 1.6e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 9 VLEEVY 14
 DB 3 VVVEVY 8

RESULT 5
 MIFH_TRISP STANDARD; PRT; 19 AA.
 AC P81529:
 DT 15-JUL-1989 (Rel. 38, Created)
 DT 15-JUL-1989 (Rel. 38, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE L-DOPACHROME-METHYL ESTER TAUOMERASE (MACROPHAGE MIGRATION INHIBITORY
 DE FACTOR HOMOLOG) (FRAGMENT).
 OS Trichinella spiralis.
 OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
 OC Trichocephalina; Trichinelloidea; Trichinelidae; Trichinella.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=99013685; PubMed=9794786;
 RA Pennock J.L., Behnke J.M., Belkile O.D., Devaney E., Grencis R.K.,
 RA Isaac R.E., Joshua G.W.P., Selkirk M.E., Zhang Y., Meyer D.J.;
 RT "Rapid purification and characterization of L-dopamine-methyl ester
 RT tautomerase (macrophage migration inhibitory factor) from Trichinella
 RT spiralis, Trichuris muris and Brugia pahangi.";
 RL Biochem. J. 335:495-498(1998).
 CC -1- FUNCTION: TAUOMERIZATION OF THE METHYL ESTER OF L-DOPACHROME.
 CC -1- INDUCTION: INHIBITED BY FREE FATTY ACIDS AND HAEMATIN.
 CC -1- SIMILARITY: BELONGS TO THE MIF FAMILY.
 CC INTERPRO: IPR001398; -;
 DR PROSITE: PS01158; MIF; PARTIAL.
 DR CYCLOKINE.
 FT NON_TER
 FT NON_TER 19
 SQ SEQUENCE 19 AA; 2106 MW; F7CA05F112A628D CRC64;

Query Match 20.2%; Score 20; DB 1; Length 19;
 Best Local Similarity 60.0%; Pred. No. 2e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 PVFVL 10
 DB 1 PIFVL 5

RESULT 6
 AMP_FUSNU STANDARD; PRT; 20 AA.
 ID AMP_FUSNU
 AC P81207;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE AMINOPEPTIDASE (EC 3.4.11.-) (AP) (FRAGMENT).
 OS Fusbacterium nucleatum.
 OC Bacteria; Fusbacteria; Fusbacterium.
 RN [1]
 RP SEQUENCE.
 RC STRAIN-ATCC 10953;
 RX MEDLINE-98361026; PubMed-9695913;
 RA Rogers A.H., Gunadi A., Gully N.J., Zilm P.S.;
 RT "An aminopeptidase nutritionally important to Fusbacterium
 nucleatum.";
 RL Microbiology 144:1807-1813(1998).
 CC -1- FUNCTION: CLEAVES A WIDE RANGE OF DIPEPTIDES AND TRIPEPTIDES.
 CC -1- COFACTOR: COBALT.
 CC -1- SUBCELLULAR LOCATION: CELL ENVELOPE-ASSOCIATED.
 CC Hydrolase; Aminopeptidase; Metalloprotease; Cobalt.
 KW NON_TER 20 20
 FT SEQUENCE 20 AA; 2585 MW; 97095B948262C71B CRC64;
 SQ

Query Match 20.2%; Score 20; DB 1; Length 20;
 Best Local Similarity 66.7%; Pred. No. 2.1e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 13 YVTLK 18
 DB 6 YVDLKE 11

RESULT 7
 MDH_THEAL
 ID MDH_THEAL STANDARD; PRT; 23 AA.
 AC P33163;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT).
 GN MDH.
 OS Thermoleophilum album.
 OC Bacteria; Green non-sulfur bacteria; Thermoleophilum.
 RN [1]
 RP SEQUENCE.
 RC STRAIN-NM;
 RA Novotny J.F., Jr., Perry J.J.;
 RT "Characterization of the malate dehydrogenase from Thermoleophilum
 album NM.";
 RL Arch. Microbiol. 154:304-307(1990).
 CC -1- CATALYTIC ACTIVITY: L-MALATE + NAD(+) -> OXALOACETATE + NADH.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
 DR PIR: A60689; A60689.
 DR INTERPRO: IPR001252;
 DR PROSITE: PS00068; MDH: PARTIAL.
 KW Oxidoreductase; Tricarboxylic acid cycle; NAD.
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2344 MW; 9848F8218EFAF125 CRC64;

Query Match 20.2%; Score 20; DB 1; Length 23;
 Best Local Similarity 60.0%; Pred. No. 2.5e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 11 LEVVT 15
 DB 17 IEXLT 21

RESULT 8
 PSAL_SPIOL
 ID PSAL_SPIOL STANDARD; PRT; 16 AA.
 AC P17230;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PHOTOSYSTEM I REACTION CENTRE SUBUNIT IX (PSI-I) (FRAGMENT).
 GN PSAL.
 OS Spinacia oleracea (Spinach).
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
 OC Caryophyllales; Chenopodiaceae; Spinacia.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-90242987; PubMed-2185953;
 RA Ikeuchi M., Hirano A., Hiyama T., Inoue Y.;
 RT "Polypeptide composition of higher plant photosystem I complex.
 RT Identification of psal, psal and psal gene products.";
 RL FEBS Lett. 263:274-278(1990).
 CC -1- FUNCTION: MAY HELP IN THE ORGANIZATION OF THE PSAL AND PSAL
 SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PSAL FAMILY.
 CC PIR: S09732; S09732.
 DR Chlroplast; Photosystem I; Photosynthesis; Transmembrane.
 KW TRANSMEM 6 >16
 FT NON_TER 16 16
 FT SEQUENCE 16 AA; 1851 MW; A1B8DA6E26E7D110 CRC64;
 SQ

Query Match 19.7%; Score 19.5; DB 1; Length 16;
 Best Local Similarity 62.5%; Pred. No. 2.1e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 1 TY-ETAPV 7
 DB 6 TYLSAVP 13

RESULT 9
 PA22_BOTSC
 ID PA22_BOTSC STANDARD; PRT; 25 AA.
 AC P80963;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE PHOSPHOLIPASE A2 HOMOLOG. MYOTOXIN II (FRAGMENT).
 OS Bothriechis schlegelii (Eyelash viper).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Viperidae; Crotalinae; Bothriechis.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-VENOM;
 RX MEDLINE-97209510; PubMed-9056257;
 RA Angulo Y., Chaves E., Alape A., Rucavado A., Gutierrez J.M.,
 RA Lomonte B.;
 RT "Isolation and characterization of a myotoxic phospholipase A2 from
 the venom of the arboreal snake Bothriechis (Bothrops) schlegelii
 from Costa Rica.";
 RL Arch. Biochem. Biophys. 339:260-266(1997).
 CC -1- FUNCTION: THIS MYOTOXIC PROTEIN LACKS ENZYMATIC ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
 DR HSSP: P24605; 1CLP.
 DR INTERPRO: IPR001211;
 DR PFAM: PF000068; phoslipd; 1.
 DR PROSITE: PS00118; PA2_HIS; PARTIAL.
 DR PROSITE: PS00119; PA2_ASP; PARTIAL.
 KW Toxin; Venom; Multigene family.
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 2725 MW; C578AB09235EA02A CRC64;

Query Match 19.7%; Score 19.5; DB 1; Length 25;
 Best Local Similarity 45.5%; Pred. No. 3.3e+03;
 Matches 5; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

OY 2 YETAPVFLVE 12

Db 3 YEIGRM-ITL 12

RESULT 10
CIP1_MYTED STANDARD; PRT: 6 AA.
ID CIP1_MYTED
AC P13736;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CONTRACTION-INHIBITING PEPTIDE I (MIP I).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
OC Mytiloidea; Mytilidae; Mytilus.
RN [1]
RP SEQUENCE.
RC TISSUE-PEDAL GANGLION; MEDLINE-88240357; PubMed-3377776;
RX Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.,
RT "Structures and actions of Mytilus inhibitory peptides."
RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
CC -I- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN MUSCLES.
CC -I- SIMILARITY: TO MIP II.
DR PIR: A27696; A27696.
KW Hormone; Amidation.
FT MOD_RES
SQ SEQUENCE 6 AA: 637 MW: 72C9C68775B81000 CRC64;

Query Match 19.2%; Score 19; DB 1; Length 6;
Best Local Similarity 60.0%; Pred. No. 8.9e+04;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 APVEV 9
Db 2 SPWFV 6

RESULT 11
UL44_HUMAN STANDARD; PRT: 14 AA.
ID UL44_HUMAN
AC P30039;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF LIVER TISSUE (SPOT 32) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RN [1]
RP SEQUENCE.
RC TISSUE-LIVER;
RX MEDLINE-93162045; PubMed-1286669;
RA Hochstrasser D.F., Frutiger S., Paquet N., Batroch A., Ravier F.,
RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargha R.,
RA Appel R.D., Hughes G.J.;
RT "Human liver protein map: a reference database established by
RT microsequencing and gel comparison."
RL Electrophoresis 13:992-1001(1992).
CC -I- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN
DR PROTEIN IS: 6.2, ITS MW IS: 28 KDA.
DR SWISS-2DPAGE; P30039; HUMAN.
FT NON_TER
SQ SEQUENCE 14 AA: 1614 MW: 0F96AC07E93295D0 CRC64;

Query Match 19.2%; Score 19; DB 1; Length 14;
Best Local Similarity 50.0%; Pred. No. 2.2e+03;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 6 PVFV 9

Db 4 PIFI 7

RESULT 12
UN04_CLOPA STANDARD; PRT: 14 AA.
ID UN04_CLOPA
AC P81349;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE UNKNOWN PROTEIN CP 4 FROM 2D-PAGE (FRAGMENT).
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
RN [1]
RP SEQUENCE.
RC STRAIN-W5;
RX MEDLINE-96291870; PubMed-9629918;
RX Flengrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5."
RL Electrophoresis 19:802-806(1998).
CC -I- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 4.7, ITS MW IS: 45.7 KDA.
DR NON_TER
FT NON_TER
SQ SEQUENCE 14 AA: 1608 MW: D424A7069460EBD2 CRC64;

Query Match 19.2%; Score 19; DB 1; Length 14;
Best Local Similarity 50.0%; Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 YEIAPVEV 9
Db 2 YEIXIVV 9

RESULT 13
MIF_PIG STANDARD; PRT: 20 AA.
ID MIF_PIG
AC P80928;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MACROPHAGE MIGRATION INHIBITORY FACTOR (MIF) (GLYCOSYLATION-INHIBITING
DE FACTOR) (GIF) (FRAGMENT).
GN MIF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE.
RC RAVIERE S., Bouet F., Menez A., Galat A.;
RL Submitted (MAR-1997) to the SWISS-PROT data bank.
CC -I- FUNCTION: THE EXPRESSION OF MIF AT SITES OF INFLAMMATION SUGGEST
CC A ROLE FOR THE MEDIATOR IN REGULATING THE FUNCTION OF MACROPHAGE
CC IN HOST DEFENSE (BY SIMILARITY).
CC -I- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE MIF FAMILY.
CC HSSP: P14174; MIF.
DR INTERPRO: IPR001398;
DR PROSITE: PS01156; MIF; PARTIAL.
DR Macrophage; Inflammatory response; Cytokine.
KW Macrophage; Inflammatory response; Cytokine.
FT NON_TER
SQ SEQUENCE 20 AA: 2147 MW: 3517AF60F3012A61 CRC64;

Query Match 19.2%; Score 19; DB 1; Length 20;
Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 6 PVFVL 10


```

Db      1 PMFV 5

RESULT 14
SUBC_CANFA STANDARD; PRT; 20 AA.
AC P99507;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 38, Last annotation update)
DE PUTATIVE SUCCINYL-COA LIGASE (GDP-FORMING) BETA-CHAIN (EC 6.2.1.4)
OS (SUCCINYL-COA SYNTHETASE, BETA CHAIN) (SCS-BETA) (FRAGMENT).
OC Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE.
RC TISSUE=HEART;
RX MEDLINE=9616340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins."
RL Electrophoresis 18:2795-2802(1997).
CC -1- CATALYTIC ACTIVITY: SUCCINATE + COA + GTP -> SUCCINYL-COA + GDP +
CC OROTHOPHOSPHATE.
CC -1- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
CC ACID CYCLE.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL (BY SIMILARITY).
CC -1- SIMILARITY: TO OTHER BETA SUBUNITS OF SUCCINYL-COA SYNTHETASE,
CC OF MALATE-COA LIGASE AND TO ATP CITRATE-LYASE.
DR HSC-2DPAGE; P99507; DOG.
DR INTERPRO: IPR000303; -.
DR PS01217; SUCCINYL-COA LIG. 3; PARTIAL.
KM Ligase: Glycolysis; Tricarboxylic acid cycle; Mitochondrion.
FT NON_TER 20
FT SEQUENCE 20 AA; 2248 MW; BE8AEFD54BDAC2E CRC64;
SQ

Query Match 19.2%; Score 19; DB 1; Length 20;
Best Local Similarity 27.3%; Pred. No. 3.2e+03;
Matches 3; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 10 LLEVYTKMKR 20
Db 3 LHEYMSMELQ 13

RESULT 15
NUO7_SOLTU STANDARD; PRT; 22 AA.
AC P80730;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE 16 KDA SUBUNIT (EC 1.6.5.3)
OS Solanum tuberosum (Potato).
OC Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
RN [1]
RP SEQUENCE.
RC STRAIN=CV. BINTJE; TISSUE=TUBER;
RA Herz U., Grohmann L.;
RT Submitted (DEC-1996) to the SWISS-PROT data bank.
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC -1- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER

```

```

CC MEMBRANE.
KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT NON_TER 22
SQ SEQUENCE 22 AA; 2435 MW; D4D64C57673C2613 CRC64;

Query Match 19.2%; Score 19; DB 1; Length 22;
Best Local Similarity 66.7%; Pred. No. 3.3e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 14 VTIKKM 19
Db 16 VYIKKM 21

RESULT 16
VGLG_RABVA STANDARD; PRT; 22 AA.
ID VGLG_RABVA
AC P15199;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE SPIKE GLYCOPROTEIN PRECURSOR (FRAGMENT).
GN G.
OS Rabies virus (strain AV01).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89150295; PubMed=3147698;
RA Poch O., Torzo N., Keith G.;
RT "Sequence of the 3386 3' nucleotides of the genome of the AV01 strain
RT rabies virus: structural similarities in the protein regions involved
RT in transcription."
RL Biochimie 70:1019-1029(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL; X13357; CA31737.1; -.
DR PIR; S07817; S07817.
KM Transmembrane; Envelope protein; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN 1
FT NON_TER 22
FT SEQUENCE 22 AA; 2469 MW; BBA59CE12F3EEB85 CRC64;
SQ

Query Match 19.2%; Score 19; DB 1; Length 22;
Best Local Similarity 80.0%; Pred. No. 3.5e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 VFVLL 11
Db 7 LFVLL 11

RESULT 17
ATPY_SOLTU STANDARD; PRT; 25 AA.
ID APTY_SOLTU
AC P80497;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ATP SYNTHASE 6 KDA SUBUNIT, MITOCHONDRIAL (EC 3.6.1.34) (FRAGMENT).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;

```

OC Solanales; Solanaceae; Solanum.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-TUBER;
 RX MEDLINE-97077345; PubMed-8919912;
 RA Jansch L., Krull V., Schmitz U.K., Braun H.P.;
 RT "New insights into the composition, molecular mass and stoichiometry
 of the protein complexes of plant mitochondria.";
 RL Plant J. 9:357-368(1996).
 CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC
 COMPONENT (CF10) SUBUNIT OF THE MITOCHONDRIAL ATPASE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
 CC CF(0): Hydrogen ion transport; Mitochondrion.
 KW NON_TER 25
 FT 25
 SQ SEQUENCE 25 AA; 3297 MW; 7D5BD65935C917A4 CRC64;

Query Match 19.2%; Score 19; DB 1; Length 25;
 Best Local Similarity 50.0%; Pred. No. 4e+03;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 6 PVFVLEY 13
 ||| |
 DB 8 PVFRRRW 15

RESULT 18
 PSAL_PEA STANDARD; PRT; 13 AA.
 ID PSAL_PEA P17229;
 AC P17229;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PHOTOSYSTEM I REACTION CENTRE SUBUNIT IX (PSI-J) (FRAGMENT).
 GN PSAL.
 OS Pisum sativum (Garden pea).
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fabales; Fabaceae; Papilionoideae; Pisum.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-90242987; PubMed-2185953;
 RA Ikeuchi M., Hirano A., Hiyaama T., Inoue Y.;
 RT Polypeptide composition of higher plant photosystem I complex.
 RT Identification of psal, psal and psak gene products.";
 RL FEBS Lett. 263:274-278(1990).
 CC -1- FUNCTION: MAY HELP IN THE ORGANIZATION OF THE PSAL AND PSAL
 SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PSAL FAMILY.
 CC PIR: S09733; S09733.
 DR Chloroplast; Photosystem I; Photosynthesis; Transmembrane.
 KW TRANSMEM 6
 FT >13
 FT NON_TER 13
 FT 13
 SQ SEQUENCE 13 AA; 1516 MW; 9E2E45D11FDE3B41 CRC64;

Query Match 18.7%; Score 18.5; DB 1; Length 13;
 Best Local Similarity 62.5%; Pred. No. 2.5e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 1 TY-ETAPV 7
 ||| |
 DB 6 TYLXVAV 13

RESULT 19
 CH10_CLOPA STANDARD; PRT; 19 AA.
 ID CH10_CLOPA P81338;
 AC P81338;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE 10 KDA CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES) (CP 31) (FRAGMENT).
 GN GROES.
 OS Clostridium pasteurianum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 RN [1]
 RP SEQUENCE.
 RC STRAIN-W5;
 RX MEDLINE-98291870; PubMed-9629918;
 RA Fleggsrud R., Skjeldal L.;
 RT "Two-dimensional gel electrophoresis separation and N-terminal
 sequence analysis of proteins from Clostridium pasteurianum W5.";
 RL Electrophoresis 19:802-806(1998).
 CC -1- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
 THE ATPASE ACTIVITY OF THE LATTER (BY SIMILARITY).
 CC -1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
 (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC INTERPRO: IPR001476;
 DR PFAM: PF00166; cpn10; 1.
 DR PROSITE: PS00681; CHAPERONINS_CPN10; PARTIAL.
 KW Chaperone.
 FT NON_TER 19
 FT 19
 SQ SEQUENCE 19 AA; 2026 MW; 7D699BD414E60A60 CRC64;

Query Match 18.7%; Score 18.5; DB 1; Length 19;
 Best Local Similarity 35.3%; Pred. No. 3.7e+03;
 Matches 6; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

OY 3 ETAPVLEYTYTKM 19
 ||| |
 DB 2 KTPV--LQDNVVKRL 15

RESULT 20
 NIFS_BRAJA STANDARD; PRT; 11 AA.
 ID NIFS_BRAJA P37030;
 AC P37030;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE NIFS PROTEIN (FRAGMENT).
 GN NIFS.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bradyrhizobium group; Bradyrhizobium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ebeling S.;
 RT Submitted (DEC-1988) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: NOT KNOWN YET.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X13691; CAA31982.1;
 DR PIR: S04875; S04875.
 DR INTERPRO: IPR000192;
 DR PROSITE: PS00595; AA_TRANSFER_CLASS_5; PARTIAL.
 KW Nitrogen fixation; Pyridoxal phosphate.
 FT NON_TER 11
 FT 11
 SQ SEQUENCE 11 AA; 1308 MW; 09F0BB69476DC404 CRC64;

Query Match 18.2%; Score 18; DB 1; Length 11;
Best Local Similarity 40.0%; Pred. No. 2.5e+03;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 APFV 9
11111
DB 6 APFV 10

RESULT 21
MILT_ONCKE

ID MILT_ONCKE STANDARD; PRT; 15 AA.

AC P81037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MILTPAIN (EC 3.4.22.-) (FRAGMENT).
OS Oncorhynchus keta (Chum salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN [1]
RP SEQUENCE AND CHARACTERIZATION.
RX MEDLINE-97397031; PubMed-9253183;
RA Kawabata C., Ichishima E.;
RT Miltpain, new cysteine proteinase from the milt of chum salmon,
RT Oncorhynchus keta.";
RL Comp. Biochem. Physiol. 117B:445-452(1997).
CC -1- FUNCTION: CYSTEINE PROTEINASE THAT HYDROLYSES BASIC PROTEINS.
CC HYDROLYZE BASIC PROTEINS SUCH AS HISTONE, SALMINE AND CLUPAINE BUT
CC NOT MILK CASEIN.
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE WITH BASIC RESIDUES AT
CC P2 AND P1.
DR MEROPS: C01.093; -
KM Hydrolyase.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1730 MW; 766B7771C0F888E7 CRC64;

Query Match 18.2%; Score 18; DB 1; Length 15;
Best Local Similarity 44.4%; Pred. No. 3.5e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 6 PPFVLETV 14
1111111
DB 2 PPFVLETV 10

RESULT 22
UC17_MAIZE

ID UC17_MAIZE STANDARD; PRT; 15 AA.

AC P80623;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLOGATED COLEOPTILE (SPOT 32)
DE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN [1]
RP SEQUENCE.
RC TISSUE-COLEOPTILE;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.5, ITS MW IS: 42.7 KDA.
DR MAIZE-2DPAGE; P80623; COLEOPTILE.
DR MAIZEDB; 123949; -
FT NON_TER 1 1

FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1554 MW; COAFF15FECEEC8 CRC64;

Query Match 18.2%; Score 18; DB 1; Length 15;
Best Local Similarity 60.0%; Pred. No. 3.5e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 PPFV 10
11111
DB 5 PPFV 9

RESULT 23
ALL2_CYPDPO

ID ALL2_CYPDPO STANDARD; PRT; 18 AA.

AC P82153;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYDIASTATIN 2.
OS Cydia pomonella (Coddling moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
RN [1]
RP SEQUENCE.
RC TISSUE-LARVA;
RX MEDLINE-98054539; PubMed-9392829;
RA Duvé H., Johnson A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily";
RL Peptides 18:1301-1309(1997).
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KM Neuropeptide: Amidation.
FT MOD_RES 18 18
SQ SEQUENCE 18 AA; 2169 MW; 8E66679CC0CDF175C CRC64;

Query Match 18.2%; Score 18; DB 1; Length 18;
Best Local Similarity 42.9%; Pred. No. 4.3e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 YEAPVF 8
111111
DB 8 YKRLPVY 14

RESULT 24
AROQ_AMYME

ID AROQ_AMYME STANDARD; PRT; 20 AA.

AC P46380;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 3-DEHYDROQUINATE DEHYDRATASE (EC 4.2.1.10) (3-DEHYDROQUINASE).
DE (TYPE II DHOASE) (FRAGMENT).
GN AROQ.
OS Amycolatopsis methanolica.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Pseudonocardiales; Pseudonocardaceae;
OC Amycolatopsis.
RN [1]
RP SEQUENCE.
RC STRAIN-NCIB 11946;
RX MEDLINE-93123995; PubMed-1479361;
RA Eversink G.J.W., Hessel G.I., Vrijsloot J.W., Cogging J.R.,
RA Dijkhuizen L.;
RT "Purification and characterization of a dual function
RT 3-dehydroquinate dehydratase from Amycolatopsis methanolica";
RL J. Gen. Microbiol. 138:2449-2457(1992).
CC -1- FUNCTION: CATALYZE A TRANS-DEHYDRATION VIA AN ENOLATE
CC INTERMEDIATE. IS INVOLVED IN BOTH THE CATABOLISM OF QUINATE AND

CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS. HAS A TEMPERATURE
 CC OPTIMAL OF 76 DEGREES CELSIUS AND PH OPTIMAL OF 9.0. TYPE II
 CC ENZYMES ARE THERMOSTABLE.
 CC -1- CATALYTIC ACTIVITY: 3-DEHYDROQUINATE - 3-DEHYDROSHIKIMATE + H(2)O.
 CC -1- PATHWAY: THIRD STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
 CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -1- SUBUNIT: HOMODODECAMER.
 CC -1- SIMILARITY: BELONGS TO THE TYPE-II 3-DEHYDROQUINASE FAMILY.
 CC INTERPRO: IPR001874; -
 DR PROSITE: PS01029; DEHYDROQUINASE II; PARTIAL.
 KW Quinate metabolism; Aromatic amino acid biosynthesis; lyase.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2197 MW; C24AA183B5CFDFA CRC64;

Query Match 18.2%; Score 18; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 VFVL 10
 ||||
 DB 3 VFVL 6

RESULT 25
 CPXX_RHQRH
 ID CPXX_RHQRH STANDARD; PRT; 20 AA.
 AC P31718;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE CYTOCHROME P450-RRI (EC 1.14.-.-) (P450-RRI) (FRAGMENT).
 OS Rhodococcus rhodochrous.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
 RN [1]
 RP SEQUENCE.
 RC STRAIN-116;
 RX MEDLINE-93238689; PubMed-8477696;
 RA Ellis L.D.; Karlson U.; Timmis K.N.;
 RT "Purification and characterization of cytochrome P450RRI from
 RT Rhodococcus rhodochrous";
 RL Eur. J. Biochem. 213:211-216(1993)
 CC -1- FUNCTION: P450-RRI CATALYZES THE O-DEALKYLATION OF 2-ETHOXYPHENOL
 CC AND 2-METHOXYPHENOL TO PRODUCE CATECHOL. THE CYTOCHROME BINDS
 CC OTHER ORTHO-SUBSTITUTED PHENOLS, INCLUDING 2-ETHOXYPHENOL,
 CC 2-METHYLPHENOL AND 2-CHLOROPHENOL.
 CC -1- INDUCTION: BY 2-ETHOXYPHENOL.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC INTERPRO: IPR001128; -
 DR PROSITE: PS00086; CYTOCHROME_P450; PARTIAL.
 KW Oxidoreductase; Monooxygenase; Electron transport; Heme.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2428 MW; BCA81586D8DD1496 CRC64;

Query Match 18.2%; Score 18; DB 1; Length 20;
 Best Local Similarity 22.2%; Pred. No. 4.7e+03;
 Matches 2; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 11 LEVYTLK 19
 |::|:::
 DB 7 LDETMEEL 15

Search completed: February 5, 2001, 10:56:03
 Job time: 503 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:49:31 ; Search time 58.67 Seconds
(without alignments)
23.147 Million cell updates/sec

Title: US-08-981-824-6

Perfect score: 99

Sequence: 1 TYEIAFVFLVLEVTLLKMR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 5001

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: PIR_66.*

1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28.3	24	2	S30373	glutathione transf
2	26.3	20	2	A57106	hull allergen Gly
3	25.3	23	2	H24735	glutathione transf
4	24.2	16	2	I57530	gene c-fms protein
5	24.2	21	2	S09517	prolamin - sorghum
6	24.2	24	2	S33237	glutathione transf
7	23.2	14	2	PH1615	Ig H chain V-D-J r
8	23.2	24	2	A35187	glutathione transf
9	23.2	24	2	A20554	hemocyanin subunit
10	23.2	25	2	S18400	carboxylic acid re
11	22.2	6	2	B27696	contraction-inhibi
12	22.2	18	2	A43826	dark-type molecula
13	22.2	19	2	S74114	3-hydroxyacyl-CoA
14	22.2	19	2	T22228	hypothetical prote
15	22.2	20	2	A48367	glutaryl-CoA dehyd
16	22.2	20	2	A46644	temperature-labile
17	22.2	22	2	S13899	glycoprotein, 11.7
18	22.2	23	2	JP0058	ribosomal protein
19	22.2	6	2	A60986	N-formyl oligopept
20	21.2	12	4	JK0315	aminotransferase c
21	21.2	15	2	PQ0750	self-incompatibili
22	21.2	21	2	B48367	glutaryl-CoA dehyd
23	21.2	21	2	T07683	protease inhibitor
24	21.2	21	4	A32902	GMP reductase / g1
25	21.2	22	2	S09021	carboxylesterase (
26	21.2	23	2	S60564	homeodomain protei
27	21.2	23	2	S37491	hypothetical prote
28	21.2	23	2	A56272	conjugation induc
29	21.2	24	2	H20554	hemocyanin subunit

30	21	21.2	24	2	PH1336
31	21	21.2	24	2	I67553
32	21	21.2	25	2	E37520
33	21	21.2	25	2	B20596
34	21	21.2	25	2	F41606
35	21	21.2	25	2	S78343
36	20.5	20.7	22	2	S00189
37	20.5	20.7	25	2	H64710
38	20	20.2	13	2	PH1599
39	20	20.2	13	2	I77387
40	20	20.2	16	2	S24667
41	20	20.2	20	2	JP0060
42	20	20.2	20	2	JP0061
43	20	20.2	21	2	B49200
44	20	20.2	21	2	I53263
45	20	20.2	22	2	PM0048

ALIGNMENTS

RESULT 1
S30373
glutathione transferase (EC 2.5.1.18) mu (isoform PI 8.5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 08-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 03-May-1996
C:Accession: S30373; S30371
R:Awasthi, S.; Singhal, S.S.; Srivastava, S.K.; Awasthi, Y.C.
Arch. Biochem. Biophys. 301, 143-150, 1993
A:Title: Purification and characterization of glutathione S-transferase of murine ova
A:Reference number: S30369; M01D:93183007
A:Accession: S30373
A:Molecule type: protein
A:Residues: 1-24 <AW1>
A>Note: this form was identified by PI 8.5
A:Accession: S30371
A:Molecule type: protein
A:Residues: 1-24 <AW2>
A>Note: this form was identified by PI 7.9
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 28.3%; Score 28; DB 2; Length 24;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 6 PFEVLELY 13
DB 15 PIRMLLEY 22
RESULT 2
A57106
hull allergen Gly m 2 - soybean (fragment)
C:Species: Glycine max (soybean)
C:Date: 27-Oct-1995 #sequence_revision 27-Oct-1995 #text_change 25-Oct-1996
R:Codina, R.; Swanson, M.C.; Fernandez-Caldas, E.; Lockey, R.F.; Trudeau, W.L.; Rama,
submitted to the Protein Sequence Database, October 1995
A:Description: Partial sequence of the Gly m 2 allergen from soybean hull.
A:Reference number: A57106
A:Accession: A57106
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <COD>
A:Experimental source: seed episperma
C:Keywords: seed

Query Match 26.3%; Score 26; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 TYEIAVPE 8
| : | : |
Db 11 TYRLGPXF 18

RESULT 3

H24735
glutathione transferase (EC 2.5.1.18) MIII - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change 18-Jun-1993
C:Accession: H24735
R:Manerlyk, B.; Allin, P.; Guttenberg, C.; Jenson, H.; Tahir, M.K.; Warholm, M.; Joriva
Proc. Natl. Acad. Sci. U.S.A. 82, 7202-7206, 1985
A:Title: Identification of three classes of cytosolic glutathione transferase common to
A:Reference number: A24735; MUID:86042634
A:Accession: H24735
A:Molecule type: protein
A:Residues: 1-23 <MAN>
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 25.3%; Score 25; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 8.6e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 6 PVFVLEY 13
| : | : | : |
Db 15 PIRMLLOY 22

RESULT 4

gene c-fms protein - mouse (fragment)
I57530
C:Species: Mus sp. (mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-May-1997
C:Accession: I57530
R:Yue, X.; Favot, P.; Dunn, T.L.; Cassidy, A.I.; Hume, D.A.
Mol. Cell. Biol. 13, 3191-3201, 1993
A:Title: Expression of mRNA encoding the macrophage colony-stimulating factor receptor
A:Reference number: I57530; MUID:93268269
A:Accession: I57530
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:S62219; NID:q386001
C:Genetics:
A:Gene: c-fms
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;

Query Match 24.2%; Score 24; DB 2; Length 16;
Best Local Similarity 55.6%; Pred. No. 8.8e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 EIAPEVLL 11
| : | : | : |
Db 2 ELGPVLL 10

RESULT 5

S09517
prolamin - sorghum (fragment)
C:Species: Sorghum bicolor (sorghum)
A:Variety: P721 lines
C:Date: 21-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 09-May-1997
C:Accession: S09517
R:Bletz, U.A.
Biochem. Genet. 20, 1039-1053, 1982
A:Title: Cereal prolamin evolution and homology revealed by sequence analysis.
A:Reference number: S09165; MUID:83153005
A:Accession: S09517

A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-21 <BIE>
C:Superfamily: zeln

Query Match 24.2%; Score 24; DB 2; Length 21;
Best Local Similarity 18.2%; Pred. No. 1.2e+03;
Matches 2; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 4 IAPVLEYV 14
| : | : | : | : | : |
Db 8 LAPAIATQFL 18

RESULT 6

S33237
glutathione transferase (EC 2.5.1.18) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 07-May-1999
C:Accession: S33237
R:Ben-Arie, N.; Khen, M.; Lancet, D.
Biochem. J. 292, 379-384, 1993
A:Title: Glutathione S-transferases in rat olfactory epithelium: purification, molecu
A:Reference number: S33237; MUID:93277499
A:Accession: S33237
A:Molecule type: protein
A:Residues: 1-24 <BEN>
A:Note: 18-ile, 21-Glu, 23-Thr and 23-Glu were also found
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 24.2%; Score 24; DB 2; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 PVFVLEYV 15
| : | : | : | : |
Db 15 PIRMLLOY 24

RESULT 7

PH1615
I9 H chain V-D-J region (clone B-less 22) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1615
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less m
A:Reference number: PH1615; MUID:93501609
A:Accession: PH1615
A:Molecule type: DNA
A:Residues: 1-14 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match 23.2%; Score 23; DB 2; Length 14;
Best Local Similarity 27.3%; Pred. No. 1.1e+03;
Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 3 EIAPEVLEY 13
| : | : | : | : |
Db 4 ESNPIYAMDY 14

RESULT 8

A33187
glutathione transferase (EC 2.5.1.18) Mu 8.3 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 23-Jun-1993

C:Accession: A35187
 R:Tsushima, S.; Maki, T.; Sato, K.
 J. Biol. Chem. 265, 7150-7157, 1990
 A:Title: Purification and characterization of glutathione transferases with an activity
 A:Reference number: A35187; MWID:90237002
 C:Accession: A35187
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-24 <RSU>
 C:Superfamily: glutathione transferase
 C:Keywords: transferase

Query Match 23.2%; Score 23; DB 2; Length 24;
 Best Local Similarity 57.1%; Pred. No. 2e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 VFVLEY 13
 : : |||||
 DB 16 IXLLLEY 22

RESULT 9
 A20554
 hemocyanin subunit I - Atlantic horseshoe crab (fragment)
 C:Species: Limulus polyphemus (Atlantic horseshoe crab)
 C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 12-Apr-1995
 C:Accession: A20554
 R:Lamy, J.; Lamy, J.; Sizaret, P.Y.; Billiard, P.; Jolles, P.; Jolles, J.; Feldmann, R.
 Biochemistry 22, 5573-5583, 1983
 A:Title: Quaternary structure of Limulus polyphemus hemocyanin.
 A:Reference number: A90478
 A:Accession: A20554
 A:Molecule type: protein
 A:Residues: 1-24 <IAM>
 C:Comment: Limulus polyphemus hemocyanin is an association of eight different subunits
 C:Superfamily: hemocyanin

Query Match 23.2%; Score 23; DB 2; Length 24;
 Best Local Similarity 36.4%; Pred. No. 2e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 APVLEVY 15
 : : |||||
 DB 7 ASIALFEHLT 17

RESULT 10
 S18400
 carboxylic acid reductase - Clostridium formicoaceticum
 C:Species: Clostridium formicoaceticum
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
 C:Accession: S18400
 R:White, H.; Felcht, R.; Huber, C.; Lotspeich, F.; Simon, H.
 Biol. Chem. Hoppe-Seyler 372, 999-1005, 1991
 A:Title: Purification and some properties of the tungsten-containing carboxylic acid red
 A:Reference number: S18400; MWID:92172285
 A:Accession: S18400
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-25 <WHI>

Query Match 23.2%; Score 23; DB 2; Length 25;
 Best Local Similarity 40.0%; Pred. No. 2e+03;
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 FVLEVYTLK 17
 : : |||||
 DB 4 FIRVDMTTLK 13

RESULT 11
 B27696
 contraction-inhibiting peptide II - blue mussel
 C:Species: Mytilus edulis (blue mussel)
 C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 08-Dec-1995
 C:Accession: B27696
 R:Hirata, T.; Kubota, I.; Iwasawa, N.; Takabatake, I.; Ikeda, T.; Muneoka, Y.
 Biochem. Biophys. Res. Commun. 152, 1376-1382, 1988
 A:Title: Structures and actions of Mytilus inhibitory peptides.
 A:Reference number: A90142; MWID:88240357
 A:Accession: B27696
 A:Molecule type: protein
 A:Residues: 1-6 <HIR>
 C:Keywords: amidated carboxyl end
 F:6/Modified site: amidated carboxyl end (Val) #status experimental

Query Match 22.2%; Score 22; DB 2; Length 6;
 Best Local Similarity 80.0%; Pred. No. 2e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 APVEV 9
 : : |||||
 DB 2 APMEV 6

RESULT 12
 A43826
 dnk-type molecular chaperone - Ajellomyces capsulata (fragment)
 N:Alternate names: 80k heat shock protein 70 homolog
 C:Species: Ajellomyces capsulata, Histoplasma capsulatum
 C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 17-Mar-1999
 C:Accession: A43826
 R:Gomez, F.J.; Gomez, A.M.; Deepe Jr., G.S.
 Infect. Immun. 60, 2565-2571, 1992
 A:Title: An 80-kilodalton antigen from Histoplasma capsulatum that has homology to he
 A:Reference number: A43826; MWID:92307875
 A:Accession: A43826
 A:Molecule type: protein
 A:Residues: 1-18 <GOM>
 A:Experimental source: extract of cell wall and cell membrane
 C:Function: involved in protein folding and assembling/disassembling of protein co
 A:Description: involved in protein folding and assembling/disassembling of protein co
 C:Superfamily: heat shock protein 70
 C:Keywords: ATP; heat shock; molecular chaperone; stress-induced protein

Query Match 22.2%; Score 22; DB 2; Length 18;
 Best Local Similarity 37.5%; Pred. No. 2.2e+03;
 Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TYELAPVF 8
 : : |||||
 DB 11 TYSVGIRF 18

RESULT 13
 S74114
 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35) - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
 C:Accession: S74114
 R:Dienaiide-Noubhani, M.; Novikov, D.; Baumgart, E.; Vanhooren, J.C.T.; Fransen, M.; G
 Eur. J. Biochem. 240, 660-666, 1996
 A:Title: Further characterization of the peroxisomal 3-hydroxyacyl-CoA dehydrogenase
 e acids di- and tri-hydroxypropanoic acids are metabolized by separate multifunctio
 A:Reference number: S74113; MWID:97008958
 A:Accession: S74114
 A:Molecule type: protein
 A:Residues: 1-19 <DIE>
 A:Experimental source: liver
 C:Superfamily: 3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydrogenase homol
 C:Keywords: fatty acid beta-oxidation; mitochondrion; NAD; oxidoreductase

Query Match 22.2% Score 22; DB 2; Length 19;
 Best Local Similarity 37.5%; Pred. No. 2.3e+03;
 Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 9 VLEEVTL 16
 :|::|:
 Db 4 ILIKHTV 11

RESULT 14

122228
 hypochetrical protein F456.5 - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T22228

R:McMurray, A.
 submitted to the EMBL Data Library, November 1995

A:Reference number: Z19534

A:Accession: T22228

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-19 <MIL>

A:Cross-references: EMBL:Z68117; PIDN:CA92179.1; GSPDB:GN00028; CESP:F456.5

A:Experimental source: clone F456

C:Genetics:

A:Gene: CESP:F456.5

A:Map position: X

Query Match 22.2% Score 22; DB 2; Length 19;
 Best Local Similarity 44.4%; Pred. No. 2.3e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 IAPVVL 12
 :|::|:
 Db 11 ICPVYAVE 19

RESULT 15

A48367
 glutaryl-CoA dehydrogenase (EC 1.3.99.7) - Pseudomonas sp. (fragment)

C:Species: Pseudomonas sp.

C>Date: 01-Dec-1993 #sequence_revision 29-Sep-1999 #text_change 29-Sep-1999

C:Accession: A48367

R:Hartel, U.; Eckel, E.; Koch, J.; Fuchs, G.; Linder, D.; Buckel, W.

Arch. Microbiol. 159, 174-181, 1993

A:Title: Purification of glutaryl-CoA dehydrogenase from Pseudomonas sp., an enzyme invc

A:Reference number: A48367; MUID:93175588

A:Accession: A48367

A:Molecule type: protein

A:Residues: 1-20 <HAR>

A:Cross-references: PIDN:AAB25430.1; PID:9265706

A:Experimental source: strain KB 740

A>Note: sequence extracted from NCBI backbone (NCBIP:125304)

C:Keywords: oxidoreductase

Query Match 22.2% Score 22; DB 2; Length 20;
 Best Local Similarity 26.7%; Pred. No. 2.4e+03;
 Matches 4; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 1 TYEAPVLEVT 15
 :|::|:
 Db 6 TFDADPYLXDLT 20

RESULT 16

A46644
 temperature-labile cholesteryl ester hydrolase - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995

C:Accession: A46644
 R:Wee, S.; Grogan, W.M.
 J. Biol. Chem. 268, 8158-8163, 1993
 A:Title: Testicular temperature-labile cholesteryl ester hydrolase. Relationship to 1
 divalent cations.
 A:Reference number: A46644; MUID:93216794

A:Accession: A46644
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-20 <MEES>
 A:Experimental source: testis
 A>Note: sequence extracted from NCBI backbone (NCBIP:129155)

Query Match 22.2% Score 22; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 PVFVLE 13
 :|::|:
 Db 9 PAAVLE 16

RESULT 17

S13899
 glycoprotein, 11.7S - frog (Lepidobatrachus laevis)

C:Species: Lepidobatrachus laevis

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997

C:Accession: S13899

R:Carroll Jr., E.J.; Wei, S.H.; Nagel, G.M.

Arch. Biochem. Biophys. 284, 346-351, 1991

A:Title: Purification, physicochemical characterization, and immunohistochemical loca

A:Reference number: S13899; MUID:91112831

A:Accession: S13899

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-22 <ARC>

Query Match 22.2% Score 22; DB 2; Length 22;
 Best Local Similarity 66.7%; Pred. No. 2.7e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 12 EYVILK 17
 :|::|:
 Db 17 DYVILK 22

RESULT 18

JP0058
 ribosomal protein L30 - Bacillus cycloheptanicus (fragment)

C:Species: Bacillus cycloheptanicus

C>Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 28-Oct-1994

C:Accession: JP0058

R:Ochi, K.

submitted to JIPID, February 1994

A:Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal

A:Reference number: JP0042

A:Accession: JP0058

A:Molecule type: protein

A:Residues: 1-23 <OCH>

C:Keywords: protein biosynthesis; ribosome

Query Match 22.2% Score 22; DB 2; Length 23;
 Best Local Similarity 80.0%; Pred. No. 2.8e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 VTLLK 18
 :|::|:
 Db 6 ITLLK 10

RESULT 19

A60986
N-formyl oligopeptide - Escherichia coli (fragment)
C:Species: Escherichia coli
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993
C:Accession: A60986
R:Broom, M.F.; Mellor, D.M.; Chadwick, V.S.
Experientia 45, 1097-1099, 1989
A:Title: Purification and amino acid sequencing of naturally occurring N-formyl-methionyl
A:Reference number: A60986, MWID:90092408
A:Accession: A60986
A:Molecule type: protein
A:Residues: 1-6

C:Comment: This hexapeptide was the longest of several N-formyl oligopeptides reported.
F:1/Modified site: N-formylmethionine #status experimental

Query Match 21.2%; Score 21; DB 2; Length 6;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 VFVLL 11
DB 2 VFLL 6

RESULT 20

JX0315
aminotransferase chimera DY18 - synthetic (fragment)
C:Species: Synthetic
C:Date: 28-May-1999 #sequence_revision 28-May-1999 #text_change 28-May-1999
C:Accession: JX0315
R: Miyazawa, K.; Kawaguchi, S.; Okamoto, A.; Kato, R.; Ogawa, T.; Kuramitsu, S.
J. Biochem. 115, 568-577, 1994
A:Title: Construction of aminotransferase chimeras and analysis of their substrate spec
A:Reference number: JX0315, MWID:94334304
A:Accession: JX0315
A:Molecule type: DNA
A:Residues: 1-12 <MY>
C:Comment: This enzyme is a chimeric enzyme of Escherichia coli aspartate aminotransfer
C:Keywords: aminotransferase

Query Match 21.2%; Score 21; DB 4; Length 12;
Best Local Similarity 42.9%; Pred. No. 2.1e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 PFEVLE 12
DB 5 PILLME 11

RESULT 21

P00750
self-incompatibility protein S3 - potato (fragment)
C:Species: Solanum tuberosum (potato)
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 23-Mar-1995
C:Accession: P00750
R: Kirch, H.H.; Uhlid, H.; Lottspeich, F.; Salamini, F.; Thompson, R.D.
Theor. Appl. Genet. 78, 581-588, 1989
A:Title: Characterization of proteins associated with self-incompatibility in Solanum tu
A:Reference number: P00749
A:Accession: P00750
A:Molecule type: protein
A:Residues: 1-15 <KIR>
A:Experimental source: style
C:Keywords: glycoprotein

Query Match 21.2%; Score 21; DB 2; Length 15;
Best Local Similarity 62.5%; Pred. No. 2.7e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 FVLEEVY 15
DB 2 FELLEVS 9

RESULT 22

B48367
glutaryl-CoA dehydrogenase (EC 1.3.99.7) - Pseudomonas sp. (fragment)
C:Species: Pseudomonas sp.
C:Date: 29-Sep-1999 #sequence_revision 29-Sep-1999 #text_change 29-Sep-1999
C:Accession: B48367
R: Hartel, U.; Eckel, E.; Koch, J.; Fuchs, G.; Linder, D.; Buckel, W.
Arch. Microbiol. 159, 174-181, 1993
A:Title: Purification of glutaryl-CoA dehydrogenase from Pseudomonas sp., an enzyme 1
A:Reference number: A48367, MWID:93175988
A:Accession: B48367
A:Molecule type: protein
A:Residues: 1-21 <HAR>
A:Cross-references: PIDN: AAB25431.1; PID: g265707
A:Experimental source: strain K 172
A:Note: sequence extracted from NCBI backbone (NCBIP:125302)
C:Keywords: oxidoreductase

Query Match 21.2%; Score 21; DB 2; Length 21;
Best Local Similarity 33.3%; Pred. No. 3.7e+03;
Matches 5; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 TYELAPVLEEVY 15
DB 6 TFDWADPILLXOLT 20

RESULT 23

T07683
proteinnase inhibitor I homolog - potato (fragment)
C:Species: Solanum tuberosum (potato)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T07683
R: Lee, J.S.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24965
A:Accession: T07683
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-21 <LEE>
A:Cross-references: EMBL: D12819; NID: g21584; PID: g21585
A:Experimental source: cv. Russet Burbank
C:Keywords: proteinase inhibitor

Query Match 21.2%; Score 21; DB 2; Length 21;
Best Local Similarity 45.5%; Pred. No. 3.7e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 VFVLEEVYTK 17
DB 10 VFVLEKYSFK 20

RESULT 24

A32902
GMP reductase / glucose-6-phosphate 1-dehydrogenase fusion protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1998 #sequence_revision 10-Sep-1998 #text_change 10-Sep-1998
C:Accession: A32902
R: Kanno, H.; Huang, I.Y.; Kan, Y.W.; Yoshida, A.
Cell 58, 595-606, 1989
A:Title: Two structural genes on different chromosomes are required for encoding the
A:Reference number: A32902; MWID:89336791
A:Accession: A32902
A:Molecule type: protein

A:Residues: 1-21 <KAN>
R:Yoshida, A.; Kan, Y.W.
Cell 62, 11-12, 1990
A:Title: Origin of "fused" glucose-6-phosphate dehydrogenase.
A:Reference number: A55177
A:Contents: annotation
C:Comment: This peptide was sequenced from a fragment of a chimeric protein produced from was later shown to be an artifact of purification and not a natural component of eryth
C:Keywords: fusion protein
F:1-13/Region: GMP reductase
F:14-21/Region: glucose-6-phosphate 1-dehydrogenase

Query Match 21.2%; Score 21; DB:4; Length 21;
Best Local Similarity 44.4%; Pred. No. 3.7e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TYEIAPVFV 9
|||:
Db 2 TYSGIPITV 10

RESULT 25

S09021
carboxylesterase (EC 3.1.1.1) R11, microsomal - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Aug-1998
C:Accession: S09021
R:Hosokawa, M.; Maki, T.; Satoh, T.
Arch. Biochem. Biophys. 277, 219-227, 1990
A:Title: Characterization of molecular species of liver microsomal carboxylesterases of
A:Reference number: S09021; MUID:90179180
A:Accession: S09021
A:Molecule type: protein
A:Residues: 1-22 <HOS>
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase

Query Match 21.2%; Score 21; DB:2; Length 22;
Best Local Similarity 62.5%; Pred. No. 3.9e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 VLEEVYTL 16
|||:
Db 15 VLGRYVSL 22

Search completed: February 5, 2001, 10:49:32
Job time: 745 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:47:34 ; Search time 68.99 Seconds
(without alignments)
9.913 Million cell updates/sec

Title: US-08-981-824-6
Perfect score: 99
Sequence: 1 TYEIAPVLELVTLKKMR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSL/gcgdata/geneseq/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	100.0	20	18	W18847
2	99	100.0	20	18	W01798
3	79	79.8	20	16	R72274
4	79	79.8	20	21	Y59574
5	66	66.7	14	16	R76651
6	66	66.7	14	18	W18859
7	63	63.6	15	18	W12408
8	50	50.5	15	18	W12407
9	49	49.5	14	16	R76652
10	49	49.5	14	18	W18860
11	45	45.5	20	16	R72375
12	45	45.5	20	21	Y59575

13	36	36.4	14	16	R76659	Peptide derived fr
14	36	36.4	14	18	W18867	65 kD Glutamic aci
15	31	31.3	10	18	W14898	Phakellistatin 9.
16	31	31.3	20	4	P30120	Sequence of peptid
17	31	31.3	20	4	P30121	Sequence of peptid
18	31	31.3	20	4	P30321	Sequence of peptid
19	30	30.3	10	18	W14897	Phakellistatin 8.
20	30	30.3	15	21	Y66492	Human Src peptide
21	30	30.3	15	21	Y66517	Human Src peptide
22	30	30.3	23	13	R22401	Modified pre seque
23	30	30.3	25	11	R06504	N-terminal seque
24	29	29.3	15	21	Y66490	Human PDGF recepto
25	29	29.3	15	21	Y66491	Human Lck peptide
26	29	29.3	15	21	Y66515	Human PDGF recepto
27	29	29.3	15	21	Y66516	Human Lck peptide
28	29	29.3	20	4	P30113	Sequence of peptid
29	29	29.3	20	4	P30118	Sequence of peptid
30	29	29.3	20	4	P30119	Sequence of peptid
31	29	29.3	20	4	P30320	Sequence of peptid
32	29	29.3	20	4	P30010	Sequence of fragme
33	28	28.3	16	17	W07285	Amphiphilic antipl
34	28	28.3	16	17	Y04725	Sindbis virus PE2
35	28	28.3	25	20	W90171	Tribelin/thrombin-b
36	27	27.3	8	18	W43400	C-reactive protein
37	27	27.3	10	14	R39584	Diuretic hormone b
38	27	27.3	15	17	W07274	Amphiphilic antipl
39	27	27.3	15	19	W48362	Cyclic peptide 1 w
40	27	27.3	15	19	W48371	Cyclic peptide 10
41	27	27.3	15	21	Y66493	Human Abi peptide
42	27	27.3	15	21	Y66518	Human Abi peptide
43	27	27.3	18	12	R10600	Peptide with motil
44	27	27.3	20	14	R41484	Peptide cbl derive
45	27	27.3	20	17	R95382	Residues 141-160 o

ALIGNMENTS

RESULT 1	
W18847	
ID	W18847 standard; peptide; 20 AA.
AC	W18847;
DT	05-JAN-1998 (first entry)
XX	
DE	65 kD Glutamic acid decarboxylase peptide fragment VI.
XX	
KW	GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
KW	Insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
KW	predisposition; autoimmune; tumour; rheumatoid arthritis;
KW	multiple sclerosis.
XX	
OS	Synthetic.
XX	
PN	DE19526561-A1.
XX	
PD	23-JAN-1997.
XX	
PF	20-JUL-1995; 95DE-1026561.
XX	
PR	20-JUL-1995; 95DE-1026561.
XX	
PA	(BOEF) BOEHRINGER MANNHEIM GMBH.
XX	
PI	Donle F, Endl J, Ganz M, Jung G, Kientsch-Engel R;
PI	Pozzilli P, Stahl P;
DR	WPI; 1997-088254/09.
XX	
PT	Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
PT	Involving intradermal admn. of auto-reactive substances
XX	

PS Claim 11; Page 10; 12pp; German.
 XX
 CC W18842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predisposition to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 XX
 SQ Sequence 20 AA:

Query Match 100.0%; Score 99; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.8e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYEIAPVFVLEVTLKKMR 20
 |||||
 Db 1 tyeiapvfvllleyvllkkmr 20

RESULT 2
 W01798
 ID W01798 standard; peptide; 20 AA.
 XX
 AC W01798;
 XX
 DT 15-OCT-1997 (first entry)
 XX
 DE Human 65 kD glutamine decarboxylase peptide.
 XX
 XX Human: glutamine decarboxylase; GAD; diagnosis; predisposition;
 KW tumour; immunological; disease; autoimmune; diabetes; reagent;
 KW determination; T cell; subpopulation; medicament; treatment;
 KW prevention; production; antigen; immunogen; tolerogen; isolation;
 KW reinjection; inactivation.
 XX
 OS Homo sapiens.
 XX
 XX DE19525784-A1.
 PN
 XX 16-JAN-1997.
 PD
 XX 14-JUL-1995; 95DE-1025784.
 XX
 XX 14-JUL-1995; 95DE-1025784.
 PR
 XX (BOEF) BOEHRINGER MANNHEIM GMBH.
 PA
 PI Albert W, Boltard C, Endl J, Jung G, Schendel D;
 PI Stahl P, Van Enderl P;
 XX
 DR WPI: 1997-078452/08.
 XX
 XX Glutamine decarboxylase peptide(s) - for diagnosis and therapy of
 PT diabetes, etc.
 XX
 PS Claim 1; Page 12; 15pp; German.

XX The present peptide is a fragment of the human 65 kD glutamine
 CC decarboxylase (GAD), which can be used to diagnose, or diagnose a
 CC predisposition to, a tumour or immunological disease, preferably an
 CC autoimmune disease, especially diabetes. It can also be used as a
 CC reagent to determine specific T cell subpopulations, in medicaments
 CC to treat or prevent immunological diseases, preferably autoimmune
 CC diseases, especially diabetes, to produce antigens, especially
 CC immunogens or tolerogens and to isolate specific T cell
 CC subpopulations, which can be used to produce antigens or for
 CC reinjection, optionally after inactivation.

XX
 SQ Sequence 20 AA:

Query Match 100.0%; Score 99; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.8e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYEIAPVFVLEVTLKKMR 20
 |||||
 Db 1 tyeiapvfvllleyvllkkmr 20

RESULT 3
 R72274
 ID R72274 standard; Peptide; 20 AA.
 XX
 AC R72274;
 XX
 DT 13-NOV-1995 (first entry)
 XX
 DE Glutamic acid decarboxylase (GAD65) fragment.
 XX
 KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
 KW insulin-dependant diabetes mellitus; stiff man disease.
 XX
 OS Homo sapiens.
 XX
 XX W09507992-A.
 PN
 XX 23-MAR-1995.
 PD
 XX 24-AUG-1994; 94WO-US09478.
 XX
 XX 17-SEP-1993; 93US-0123859.
 PR
 XX (REBC) UNIV CALIFORNIA.
 PA
 PI Clare-Salier MJ, Erlander MG, Kaufman DL, Tobin AJ;
 XX
 DR WPI: 1995-131360/17.
 XX
 XX New polypeptide fragments of glutamic acid decarboxylase - for
 PT diagnosis and treatment of autoimmune disease, esp. insulin
 PT dependent diabetes, also related nucleic acid, vectors,
 PT antibodies, hybridoma(s) etc.
 XX
 PS Example 11; Page 76; 100pp; English.
 XX
 CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
 CC acid decarboxylase (GAD65) respectively, from which the GAD65
 CC fragments described in R72261-R72298 were derived. These fragments
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose
 CC and treat GAD-related autoimmune disorders, such as insulin
 CC dependant diabetes mellitus or stiff man disease.
 XX
 XX Sequence 20 AA:

Query Match 79.8%; Score 79; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.1e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYEIAPVFVLEVTL 16
 |||||
 Db 5 tyeiapvfvllleyvll 20

RESULT 4
 Y59574
 ID Y59574 standard; peptide; 20 AA.
 XX
 AC Y59574;

```

XX      03-APR-2000   (first entry)
DT
XX
XX      GAD65 fragment, peptide #14.
DE
XX
XX      GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
KW      insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
        therapy.
XX
XX      Homo sapiens.
OS
XX
XX      US5998366-A.
PN
XX
XX      07-DEC-1999.
PD
XX
XX      09-APR-1997;    97US-0827618.
PF
XX
XX      07-JUN-1995;    95US-0485725.
PR      21-SEP-1990;    90US-0586536.
PR      18-JUN-1991;    91US-0716909.
XX
XX      (REGC ) UNIV CALIFORNIA.
PA
XX      Tobin AJ, Kaufman DL, Erlander MG;
PI
XX      WPI; 2000-095930/08.
DR
XX
XX      Ameliorating glutamic acid decarboxylase associated autoimmune
PT      disorders such as insulin dependent diabetes mellitus and Stiff man
PS      disease -
XX
XX      Example 11; Column 42; 61pp; English.
PS
XX
XX      This sequence represents a fragment of the glutamic acid decarboxylase 65
CC      (GAD65) protein. The invention relates to a method of ameliorating GAD
CC      associated autoimmune disorders by administering a GAD65 peptide to the
CC      patient. The method can be used for ameliorating GAD associated
CC      autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
CC      and Stiff man disease. GAD65 can also be useful for screening drugs that
CC      alter GAD function, for generating monoclonal antibodies and in
CC      immunossays. GAD65 is an effective diagnostic tool for predicting IDDM
CC      and the diagnosis is quite easy. It is also possible to obtain much
CC      larger quantities of polypeptide via recombinant techniques than are
CC      available from natural sources.
XX
XX      Sequence     20 AA:
SQ
XX
XX      Query Match          79.8%; Score 79; DB 21; Length 20;
XX      Best Local Similarity 100.0%; Pred. No. 7.le-08;
XX      Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY      1 TYEIAPEVLELEYVL 16
       |||||
Db      5 tyeiapvfvllleyvl 20
XX
XX
RESULT      5
R76651
ID      R76651 standard; peptide; 14 AA.
XX
XX      R76651;
AC
XX
XX      05-MAR-1996 (first entry)
DT
XX
XX      Peptide derived from human glutamic acid decarboxylase 10.
DE
XX
XX      diabetes; T-cell subpopulation; detection; antigen production;
KW      diagnosis; autoimmune disease.
XX
XX      Homo sapiens.
OS
XX
XX      DE4418091-A1.
PN

```

[illegible]

DR WPI: 1997-088254/09.
 XX Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
 PR involving intradermal admin. of auto-reactive substances
 XX
 PS Claim 11: Fig 2; 12pp; German.
 CC W18842-70 are peptide fragments of the 65 kD human glutamic acid
 CC decarboxylase (GAD). The fragments are autoreactive substances used for
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
 CC determined by using a claimed method for diagnosis of cell-mediated
 CC diseases or a predisposition to cell-mediated diseases, which is effected
 CC by administering an autoreactive substance intradermally and establishing
 CC the diagnosis on the basis of the occurrence or lack of a positive
 CC reaction at the site of administration. The method is used for diagnosis
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
 XX
 SQ Sequence 14 AA:

Query Match 66.7%; Score 66; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTEIAPVFLLEY 13
 |||||
 DB 2 tyeiapvfvlley 14

RESULT 7
 W12408 standard; peptide: 15 AA.

AC W12408;
 XX
 DT 08-OCT-1997 (first entry)
 XX
 DE GAD65 residues 213-227.
 XX
 KW 65 kD glutamic acid decarboxylase; human; GAD65; enzyme; pancreas;
 KW neuron; central nervous system; type I diabetes; autoimmune response;
 KW T cell; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..15
 FT /note= "optionally substituted, providing at least 7
 FT residues remain wild type"

W09700891-A1.

09-JAN-1997.

24-JUN-1996; 96WO-US10790.

23-JUN-1995; 95US-0494624.

(KENN-) KENNEDY INST RHEUMATOLOGY.
 PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
 PA (SBAR-) ST BARTHOLOMEW'S HOSPITAL CENT CLINICAL.

Conlon PJ, Gaur A, Leslie RDG, Ling N, Londei M;
 WPI: 1997-087322/08.

New human glutamic acid decarboxylase peptide(s) - used for
 treatment, diagnosis and determining predisposition to diabetes and
 for ameliorating auto-immune responses.

Claim 26; Page -: 28pp; English.

CC W12403-W12413 represent fragments and analogues of the 65 kD isoform
 CC of human glutamic acid decarboxylase (GAD65) (see W12402 for full length
 CC wild type protein). GAD is an enzyme expressed in the beta cells of the
 CC pancreas, and in neurons of the central nervous system. There are two
 CC isoforms of GAD, a 67 kD isoform, and GAD65. Immunodominant regions of
 CC GAD65 have been identified in Type I diabetic patients. These GAD65
 CC fragments, and analogues, are used in the methods of the invention. The
 CC methods are for detecting or treating diabetes or a predisposition to
 CC diabetes. The peptides can also be used for ameliorating an autoimmune
 CC response in a patient. Alteration of the native peptides with selective
 CC changes of crucial residues can induce unresponsiveness or change the
 CC responsiveness of antigen-specific autoreactive T cells. The peptide
 CC analogues compete for binding to MHC and do not cause proliferation of
 CC the corresponding native peptide-specific T cells.
 XX
 SQ Sequence 15 AA:

Query Match 63.6%; Score 63; DB 18; Length 15;
 Best Local Similarity 100.0%; Pred. No. 3.8e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 FYLLEYVTLKKMR 20
 |||||
 DB 1 fyleyvtlkkmr 13

RESULT 8
 W12407 standard; peptide: 15 AA.

AC W12407;
 XX
 DT 08-OCT-1997 (first entry)
 XX
 DE GAD65 residues 201-215.
 XX

KW 65 kD glutamic acid decarboxylase; human; GAD65; enzyme; pancreas;
 KW neuron; central nervous system; type I diabetes; autoimmune response;
 KW T cell; therapy.
 XX

OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Misc-difference 1..15
 FT /note= "optionally substituted, providing at least 7
 FT residues remain wild type"

W09700891-A1.

09-JAN-1997.

24-JUN-1996; 96WO-US10790.

23-JUN-1995; 95US-0494624.

(KENN-) KENNEDY INST RHEUMATOLOGY.
 PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
 PA (SBAR-) ST BARTHOLOMEW'S HOSPITAL CENT CLINICAL.

Conlon PJ, Gaur A, Leslie RDG, Ling N, Londei M;
 WPI: 1997-087322/08.

New human glutamic acid decarboxylase peptide(s) - used for
 treatment, diagnosis and determining predisposition to diabetes and
 for ameliorating auto-immune responses.

Claim 25; Page -: 28pp; English.

CC W12403-W12413 represent fragments and analogues of the 65 kD isoform
 CC of human glutamic acid decarboxylase (GAD65) (see W12402 for full length
 CC wild type protein). GAD is an enzyme expressed in the beta cells of the

CC pancreas, and in neurons of the central nervous system. There are two
 CC isoforms of GAD, a 67 kD isoform, and GAD65. Immunodominant regions of
 CC GAD65 have been identified in Type I diabetic patients. These GAD65
 CC fragments, and analogues, are used in the methods of the invention. The
 CC methods are for detecting or treating diabetes or a predisposition to
 CC diabetes. The peptides can also be used for ameliorating an autoimmune
 CC response in a patient. Alteration of the native peptides with selective
 CC changes of crucial residues can induce unresponsiveness or change the
 CC responsiveness of antigen-specific autoreactive T cells. The peptide
 CC analogues compete for binding to MHC and do not cause proliferation of
 CC the corresponding native peptide-specific T cells.

XX Sequence 15 AA;

Query Match 50.5%; Score 50; DB 18; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0084;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVEIAPFVL 10
 |||||
 Db 6 tyeIapfvvl 15

RESULT 9

R76652 R76652 standard; peptide; 14 AA.

AC R76652;

XX 05-MAR-1996 (first entry)

XX Peptide derived from human glutamic acid decarboxylase 11.

XX diabetes; T-cell subpopulation; detection; antigen production;

KW diagnosis; autoimmune disease.

XX Homo sapiens.

PN DE4418091-A1.

PD 27-JUL-1995.

PF 24-MAY-1994; 94DE-4418091.

XX 04-FEB-1994; 94DE-4403522.

PR 20-JAN-1994; 94DE-4401629.

XX (EMBL/) ENDL J.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Albert W, Dormair K, Endl J, Jung G, Meinel E;

XX Stahl P, Schendel D;

DR WPI; 1995-264505/35.

PS Claim 1; Fig 2; 21pp; German.

XX R76642-62 are derived from human glutamic acid decarboxylase and

CC specifically react with T-cell sub-populations isolated from recently

CC diagnosed type-I diabetes. Pharmaceutical compns. contg. these

CC peptides and those shown in R7571-72, are useful for the diagnosis of

CC a disease or predispositions of immune system diseases, tumours, and

CC autoimmune diseases, including diabetes. The peptides are able to detect

CC specific T-cell subpopulations that are then used for antigen prodn.,

XX e.g. by reinjection.

XX Sequence 14 AA;

Query Match 49.5%; Score 49; DB 16; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.012;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LEYVTLKKMR 20
 |||||
 Db 1 leyvtlkkmr 10

RESULT 10

W18860 W18860 standard; peptide; 14 AA.

AC W18860;

XX 05-JAN-1998 (first entry)

XX 65 kD Glutamic acid decarboxylase peptide fragment 11.

XX GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;

KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;

KW predisposition; autoimmune; tumour; rheumatoid arthritis;

XX multiple sclerosis.

XX Synthetic.

XX DE19526561-A1.

PD 23-JAN-1997.

PF 20-JUL-1995; 95DE-1026561.

XX 20-JUL-1995; 95DE-1026561.

XX (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;

XX Pozzilli P, Stahl P;

DR WPI; 1997-088254/09.

PS Skin test for diagnosis of cell-mediated diseases, esp. diabetes -

PT involving intradermal admin. of auto-reactive substances

XX Claim 11; Fig 2; 12pp; German.

XX W18842-70 are peptide fragments of the 65 kD human glutamic acid

CC decarboxylase (GAD). The fragments are autoreactive substances used for

CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is

CC determined by using a claimed method for diagnosis of cell-mediated

CC diseases or a predisposition to cell-mediated diseases, which is effected

CC by administering an autoreactive substance intradermally and establishing

CC the diagnosis on the basis of the occurrence or lack of a positive

CC reaction at the site of administration. The method is used for diagnosis

CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases

CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.

XX Sequence 14 AA;

Query Match 49.5%; Score 49; DB 18; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.012;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LEYVTLKKMR 20

|||||

Db 1 leyvtlkkmr 10

RESULT 11

R72275

ID R72275 standard; peptide; 20 AA.

XX AC R72275;
 XX 13-NOV-1995 (first entry)
 DT XX
 DE Glutamic acid decarboxylase (GAD65) fragment.
 XX
 XX Glutamic acid decarboxylase; GAD65; autoimmune disorders;
 KW insulin-dependent diabetes mellitus; stiff man disease.
 XX
 OS Homo sapiens.
 XX
 PN W09507992-A.
 PD 23-MAR-1995.
 XX
 PF 24-AUG-1994; 94WO-US09478.
 XX
 PR 17-SEP-1993; 93US-0123859.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PI
 PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
 DR WPI; 1995-131360/17.
 XX
 PT New polypeptide fragments of glutamic acid decarboxylase - for
 PT diagnosis and treatment of autoimmune disease, esp. insulin
 PT dependent diabetes, also related nucleic acid, vectors,
 PT antibodies, hybridoma(s) etc.
 XX
 PS Example 11: Page 76; 100pp; English.
 XX
 CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
 CC acid decarboxylase (GAD65) respectively, from which the GAD65
 CC fragments described in R72261-R72298 were derived. These fragments
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose
 CC and treat GAD-related autoimmune disorders, such as insulin
 CC dependant diabetes mellitus or stiff man disease.
 CC
 SQ Sequence 20 AA;

Query Match 45.5%; Score 45; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.094; 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;
 OY 12 EYVTLKKMR 20
 DB 1 eyvtlikmr 9

RESULT 12
 ID Y59575
 XX K59575 standard; peptide: 20 AA.
 AC Y59575;
 XX
 XX 03-APR-2000 (first entry)
 DT XX
 DE GAD65 fragment, peptide #15.
 XX
 KW GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
 KW insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
 KW therapy.
 XX
 OS Homo sapiens.
 XX
 PN US5998366-A.
 PD 07-DEC-1999.
 XX
 PF 09-APR-1997; 97US-0827618.

XX 07-JUN-1995; 95US-0485725.
 PR 21-SEP-1990; 90US-0586536.
 PR 18-JUN-1991; 91US-0716909.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 PI Tobin AJ, Kaufman DL, Erlander MG;
 DR WPI; 2000-095930/08.
 XX
 PT Ameliorating glutamic acid decarboxylase associated autoimmune
 PT disorders such as insulin dependent diabetes mellitus and Stiff man
 PT disease -
 XX
 PS Example 11: Column 42; 61pp; English.
 XX
 CC This sequence represents a fragment of the glutamic acid decarboxylase 65
 CC (GAD65) protein. The invention relates to a method of ameliorating GAD
 CC associated autoimmune disorders by administering a GAD65 peptide to the
 CC patient. The method can be used for ameliorating GAD associated
 CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
 CC and Stiff man disease. GAD65 can also be useful for screening drugs that
 CC alter GAD function, for generating monoclonal antibodies and in
 CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
 CC and the diagnosis is quite easy. It is also possible to obtain much
 CC larger quantities of polypeptide via recombinant techniques than are
 CC available from natural sources.
 XX
 SQ Sequence 20 AA;

Query Match 45.5%; Score 45; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.094;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 12 EYVTLKKMR 20
 DB 1 eyvtlikmr 9

RESULT 13
 ID R76659
 XX R76659 standard; peptide; 14 AA.
 AC R76659;
 XX
 XX 05-MAR-1996 (first entry)
 DT XX
 DE Peptide derived from human glutamic acid decarboxylase 18.
 XX
 KW diabetes; T-cell subpopulation; detection; antigen production;
 KW diagnosis; autoimmune disease.
 KW
 OS Homo sapiens.
 XX
 PN DE4418091-A1.
 PD 27-JUL-1995.
 XX
 PF 24-MAY-1994; 94DE-4418091.
 XX
 PR 04-FEB-1994; 94DE-4403522.
 PR 20-JAN-1994; 94DE-4401629.
 XX
 XX (ENDL/) ENDL J.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 XX
 PI Albert W, Dormair K, Endl J, Jung G, Meinel E;
 PI Stahl P, Schendel D;
 XX
 DR WPI; 1995-264505/35.
 XX

PT Antigen-specific activated T-lymphocytes and their detection - by
PT Interaction with Inventive peptide(s) of peptide-MHC complexes;
PT useful in diagnosis of e.g. diabetes and auto-immune diseases
XX
PS Claim 1; Fig 2; 21pp; German.
XX
CC R76642-62 are derived from human glutamic acid decarboxylase and
CC specifically react with T-cell sub-populations isolated from recently
CC diagnosed Type-1 diabetics. Pharmaceutical compns. contg. these
CC peptides and those shown in R77571-72, are useful for the diagnosis of
CC a disease or predispositions of immune system diseases, tumours, and
CC autoimmune diseases, including diabetes. The peptides are able to detect
CC specific T-cell subpopulations that are then used for antigen prodn.,
CC e.g. by reinfection.
XX
SQ Sequence 14 AA;
XX
Query Match 36.4%; Score 36; DB 16; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TYETAPV 7
DB 8 lyeiapv 14
XX
RESULT 14
ID W18867 standard; peptide; 14 AA.
XX
AC W18867;
XX
DT 05-JAN-1998 (first entry)
XX
DE 65 kD Glutamic acid decarboxylase peptide fragment 18.
XX
KW GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;
KW Insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
KM Predisposition; autoimmune; tumour; rheumatoid arthritis;
KM Multiple sclerosis.
XX
OS Synthetic.
XX
PN DE19526561-A1.
XX
PD 23-JAN-1997.
XX
PF 20-JUL-1995; 95DE-1026561.
XX
PR 20-JUL-1995; 95DE-1026561.
XX
PA (BOE) BOEHRINGER MANNHEIM GMBH.
XX
PI Donie F, Endl J, Ganz M, Jung G, Klentsch-engel R;
PI Pozzilli P, Stahl P;
XX
DR WPI; 1997-088254/09.
XX
PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
PT Involving intradermal admin. of auto-reactive substances
XX
PS Claim 11; Fig 2; 12pp; German.
XX
CC W18842-70 are peptide fragments of the 65 kD human glutamic acid
CC decarboxylase (GAD). The fragments are autoreactive substances used for
CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
CC determined by using a claimed method for diagnosis of cell-mediated
CC diseases or a predisposition to cell-mediated diseases, which is effected
CC by administering an autoreactive substance intradermally and establishing
CC the diagnosis on the basis of the occurrence or lack of a positive
CC reaction at the site of administration. The method is used for diagnosis
CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases

CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
XX
SQ Sequence 14 AA;
XX
Query Match 36.4%; Score 36; DB 18; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TYETAPV 7
DB 8 lyeiapv 14
XX
RESULT 15
ID W14898 standard; peptide; 10 AA.
XX
AC W14898;
XX
DT 27-NOV-1997 (first entry)
XX
DE Phakellistatin 9.
XX
KW cyclic; decapeptide; phakellistatin; sponge; cytostatic;
KW anticancer.
XX
OS Phakellia costata.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note- "the imino group of this residue forms a
FT peptide bond with the carboxy group of Val(10),
FT i.e. the residue is a component of a cyclic
FT peptide and is not an N-terminal"
FT Modified-site 10 /note- "the carboxy group of this residue forms a
FT peptide bond with the imino group of Pro(1),
FT i.e. the residue is a component of a cyclic
FT peptide and is not a C-terminal"
XX
PN US5646246-A.
XX
PD 08-JUL-1997.
XX
PF 14-NOV-1994; 94US-0338383.
XX
PR 14-NOV-1994; 94US-0338383.
XX
PA (UYAR-) UNIV ARIZONA STATE.
XX
PI Pettit GR, Xu J;
XX
DR WPI; 1997-362994/33.
XX
PT New cyclic peptide(s) designated phakellistatin -4, -5, -6, -7, -8
PT and -9 - are isolated from the sponge Phakellia costata and may be
PT used in treatment of cancer
XX
PS Claim 7; Column 55; 29pp; English.
XX
CC The patent discloses six new cyclic peptides which have been isolated
CC from the Western Pacific Ocean sponge Phakellia costata. They are:
CC cyclo-(Pro-Thr-Pro-Phe-Ile-Phe-Ser) (designated phakellistatin 4),
CC cyclo-(Pro-Phe-Asn-Ala-Met-Ala-Ile) (designated phakellistatin 5),
CC cyclo-(Pro-Trp-Leu-Pro-Ile-Pro-Phe) (designated phakellistatin 6),
CC cyclo-(Pro-Pro-Ile-Phe-Ala-Leu-Pro-Pro-Tyr-Ile) (designated
CC phakellistatin 7), cyclo-(Pro-Pro-Ile-Phe-Val-Leu-Pro-Tyr-Ile)
CC (designated phakellistatin 8) and cyclo-(Pro-Pro-Ile-Phe-Val-Leu-Pro-
CC Pro-Tyr-Val) (designated phakellistatin 9). The cyclic peptides all
CC demonstrate in vitro p38 ED50 values of between 0.18 and 4.1 mcg/ml.
CC They also demonstrate in vitro growth inhibitory activity against
CC human cancer cell lines. They may thus be used in treatment of cancer.

XX The invention relates to inhibiting a protein kinase (I) that has Thr (or
CC a residue less bulky than Thr) at the position equivalent to Thr 106 in
CC SAPK2/p38 by treatment with a pyridyl-1-imidazole, or related inhibitor.
CC The (I) is not SAPK2/p38 or SAPK2/p38beta2. The inhibitors are used
CC therapeutically: (i) to treat extracellular matrix (ECM) deposition, to
CC encourage tissue regeneration and/or repair, tissue remodelling and
CC healing of wounds or injuries, also to reduce scar tissue formation
CC following a brain injury; (ii) to treat patients with (or at risk of
CC developing) end-stage organ failure, pathological ECM accumulation,
CC (renal) fibrosis, immunosuppression, diabetic nephropathy, and kidney
CC damage; (iii) to treat defects of bone growth or homeostasis (e.g.
CC osteoporosis), arthritis and atherosclerosis in which transforming growth
CC factor-beta (but not proinflammatory cytokines) are implicated, and (iv)
CC to treat epithelial, mesenchymal or glial cancers, solid tumors,
CC melanomas, small cell lung cancer, chronic myelogenous leukemia
CC (including purging bone marrow of affected cells), and T cell lymphoma.
CC Transgenic animals (or transfect cells) that include a protein kinase
CC in which the position equivalent to 106 is mutated are used to determine
CC the physiological role of the kinase and to identify its substrates, and
CC the protein kinase can be used to screen for drugs (or lead compounds),
CC The inhibitors are specific for the (I), i.e. they do not inhibit similar
CC protein kinases in which the position equivalent to Thr 106 is occupied
CC by Met or some other bulky residue.
XX
XX
XX Sequence 15 AA:
XX

Query Match	30.3%	Score 30;	DB 21;	Length 15;
Best Local Similarity	30.0%;	Pred. No. 34;		
Matches	3;	Conservative	6;	Mismatches 1; Indels 0; Gaps 0

Qy	6	pvevlllevt	15
	1	::: 11::	
Db	3	plyivteyrs	12

```

RESULT 22
R22401
ID      R22401 standard; Protein; 23 AA.

```

DE	Modified pre sequence for AOA.
XX	
DI	29-JUL-1992 (first entry)
DI	

KW Acylxoyacyl hydrolase; septicemia; LPS; vaccine; human.
XX
XX
OS Synthetic.
XX

XX
PD
XX

19-MAR-1992.

[illegible]

XX

PA (TEXA) UNIV OF TEXAS SYSTEM.

PI Ohara PJ, Hagen FS, Grant FJ, Munford RS;
 VV

DR WPI; "1992-114352/14.
XX

PT constructed encoding whole protein or its sub-units, and used to

PT treat gram negative sepsis.
XX

PS Claim 18; page 6 /; 9/pp; English.
XX

CC DNA encoding the sequence was ligated to the *hcmv* gene in an

expression vector. The vector was used to express the recombinant enzyme, which can be used prophylactically and therapeutically against Gram-negative septicemia in mammals. For treatment, 1 ug-10 mg AOH is given per 70 kg body wt. The recombinant enzyme is produced in higher levels than found naturally in neutrophils and is more easily purified. See also R22395-400.

Sequence 23 AA;

Query Match 30.3%; Score 30; DB 13; Length 23;

Best Local Similarity 62.5%; Pred. No. 56; Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 IAPFVLL 11
: : : : :
Db 10 vapflll 17

RESULT 23

R06504 R06504 standard; protein; 25 AA.

XX R06504;

DT 07-JAN-1991 (first entry)

XX N-terminal sequence of Mr.26000 of F.hepatoca GST (Fh26b).

XX Glutathione-S-transferase (GST); liver fluke infection vaccine; Fh26b.

XX Fasciola hepatica.

PN WO9008819-A.

PD 09-AUG-1990.

XX 31-JAN-1990; 90WO-AU00027.

PR 31-JAN-1989; 89AU-0002481.

PA (DARA-) DARATECH PTY LTD.

PI Camerli SL, Panaccio M, Wilson R, Wiffels GL, Spithill TW; Campbell NJ, Thompson C, Sexton JL;

DR WPI; 1990-260928/34.

PT Vaccine against liver fluke infection in ruminants. esp. sheep - comprises glutathione-S-transferase extracted from adult fasciola hepatica worms.

PS Disclosure; Fig 19; 44pp; English.

XX Fh26b is a GST isolated from the liver fluke F.hepatoca. Its N-terminal sequence shows very high levels of homology with the

XX Corresponding region of Schistosoma and mammalian GSTs.

CC Conservation of several key regions of sequence resulted in

CC identities of 55-77% for Fh26b.

CC See also Q05732-Q05736, R06503 and R06505-R06506.

XX Sequence 25 AA;

Query Match 30.3%; Score 30; DB 11; Length 25;

Best Local Similarity 66.7%; Pred. No. 62; Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 PVEVLETV 14
: : : : :
Db 15 pvrllleyl 23

RESULT 24

ID Y66490 standard; peptide; 15 AA.

XX Y66490;

DT 07-MAR-2000 (first entry)

XX Human PDGF receptor peptide fragment.

XX Protein kinase; inhibitor; extracellular matrix deposition; ECM; cancer; tissue regeneration; tissue remodeling; wound healing; injury; fibrosis;

XX Immunosuppression; diabetic neuropathy; and kidney damage; bone growth;

XX MAP kinase; osteoporosis; arthritis; atherosclerosis; cytokine; tumor.

OS Homo sapiens.

PN WO958128-A1.

PD 18-NOV-1999.

XX 04-MAY-1999; 99WO-GB01385.

PR 09-MAY-1998; 98GB-0009869.

XX (MEDI-) MEDICAL RES COUNCIL.

PI Cohen P, Goedert M, Evers PA;

DR WPI; 2000-038954/03.

XX Inhibiting particular protein kinases that have non-bulky residue in the ATP binding site using pyridinyl-imidazole derivative, e.g. for

XX reducing deposition of extracellular matrix -

XX Example 1; Fig 1; 93pp; English.

XX The invention relates to inhibiting a protein kinase (I) that has Thr (or

XX a residue less bulky than Thr) at the position equivalent to Thr 106 in

XX SAPK2a/p38 by treatment with a pyridinyl-imidazole, or related inhibitor.

XX The (I) is not SAPK2a/p38 or SAPK2b/p38beta2. The inhibitors are used

XX therapeutically: (i) to treat extracellular matrix (ECM) deposition, to

XX encourage tissue regeneration and/or repair, tissue remodeling and

XX healing of wounds or injuries, also to reduce scar tissue formation

XX following a brain injury; (ii) to treat patients with (or at risk of

XX developing) end-stage organ failure, pathological ECM accumulation,

XX (renal) fibrosis, immunosuppression, diabetic neuropathy, and kidney

XX damage; (iii) to treat defects of bone growth or homeostasis (e.g.

XX osteoporosis), arthritis and atherosclerosis in which transforming growth

XX factorbeta (but not proinflammatory cytokines) are implicated, and (iv)

XX to treat epithelial, mesenchymal or glial cancers, solid tumors,

XX melanomas, small cell lung cancer, chronic myelogenous leukemia

XX (including purging bone marrow of affected cells) and T cell lymphoma.

XX Transgenic animals (or transfected cells) that include a protein kinase

XX in which the position equivalent to 106 is mutated are used to determine

XX the physiological role of the kinase and to identify its substrates, and

XX the protein kinase can be used to screen for drugs (or lead compounds).

XX The inhibitors are specific for the (i), i.e. they do not inhibit similar

XX protein kinases in which the position equivalent to Thr 106 is occupied

XX by Met or some other bulky residue.

XX Sequence 15 AA;

Query Match 29.3%; Score 29; DB 21; Length 15;

Best Local Similarity 37.5%; Pred. No. 51; Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 PVEVLETV 13
: : : : :
Db 3 pvrllleyl 10

RESULT 25

ID Y66491 standard; peptide: 15 AA.

AC Y66491:

DT 07-MAR-2000 (first entry)

DE Human Lck peptide fragment.

Protein kinase; inhibitor; extracellular matrix deposition; ECM; cancer;
 tissue regeneration; tissue remodeling; wound healing; injury; fibrosis;
 immunosuppression; diabetic nephropathy; and kidney damage; bone growth;
 MAP kinase; osteoporosis; arthritis; atherosclerosis; cytokine; tumor.

OS Homo sapiens.

PN W09958128-A1.

PD 18-NOV-1999.

PF 04-MAY-1999; 99MO-GB01385.

PR 09-MAY-1998; 98GB-0009869.

PA (MED1-) MEDICAL RES COUNCIL.

PI Cohen P, Goedert M, Eyers PA;

DR WPI: 2000-038954/03.

Inhibiting particular protein kinases that have non-bulky residue in
 the ATP binding site using pyridinyl-imidazole derivative, e.g. for
 reducing deposition of extracellular matrix -

PS Example 1; Fig 1; 93pp; English.

CC The invention relates to inhibiting a protein kinase (I) that has Thr (or
 CC a residue less bulky than Thr) at the position equivalent to Thr 106 in
 CC SAPK2a/p38 by treatment with a pyridinyl-imidazole, or related inhibitor.
 CC The (I) is not SAPK2a/p38 or SAPK2b/p38beta2. The inhibitors are used
 CC therapeutically: (i) to treat extracellular matrix (ECM) deposition, to
 CC encourage tissue regeneration and/or repair, tissue remodeling and
 CC healing of wounds or injuries, also to reduce scar tissue formation
 CC following a brain injury; (ii) to treat patients with (or at risk of
 CC developing) end-stage organ failure, pathological ECM accumulation,
 CC (renal) fibrosis, immunosuppression, diabetic nephropathy, and kidney
 CC damage; (iii) to treat defects of bone growth or homeostasis (e.g.
 CC osteoporosis), arthritis and atherosclerosis in which transforming growth
 CC factorbeta (but not proinflammatory cytokines) are implicated, and (iv)
 CC to treat epithelial, mesenchymal or glial cancers, solid tumors,
 CC melanomas, small cell lung cancer, chronic myelogenous leukemia
 CC (including purging bone marrow of affected cells) and T cell lymphoma.
 CC Transgenic animals (or transfected cells) that include a protein kinase
 CC in which the position equivalent to 106 is mutated are used to determine
 CC the physiological role of the kinase and to identify its substrates, and
 CC the protein kinase can be used to screen for drugs (or lead compounds).
 CC The inhibitors are specific for the (I); i.e. they do not inhibit similar
 CC protein kinases in which the position equivalent to Thr 106 is occupied
 CC by Met or some other bulky residue.

SQ Sequence 15 AA;

Query Match

29.3%; Score 29; DB 21; Length 15;

Best Local Similarity 37.5%; Pred. No. 51;

Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 PVEVLLEY 13

DB 3 pylvlley 10

Search completed: February 5, 2001, 10:47:34
 Job time: 628 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:55:28 ; Search time 93.77 Seconds
(without alignments)
24.999 Million cell updates/sec

Title: US-08-981-824-7

Sequence: 1 FFRVYISNPATRHQDIDFLI 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.podent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	26.2	14	2 09KEL4	09KEL4 bacillus ha
2	27	26.2	15	5 09TWT4	09TWT4 lumbricus t
3	26	25.2	17	6 062645	062645 sagittus oe
4	26	25.2	18	6 046665	046665 macropus ro
5	26	25.2	19	11 09JIC5	09JIC5 mus musculu
6	26	25.2	24	12 084272	084272 human papil
7	25	24.3	18	10 010722	010722 engelmannia
8	25	24.3	19	10 040183	040183 lemna gibba
9	25	24.3	21	10 040181	040181 lemna gibba
10	24	23.3	8	11 062527	062527 mus. spretus
11	24	23.3	20	2 09R4A6	09R4A6 hydrogencba
12	24	23.3	20	12 069381	069381 herpes simp
13	24	23.3	21	12 093G25	093G25 tt virus. o
14	24	23.3	22	12 068988	068988 human herpe
15	24	23.3	22	12 090G68	090G68 human immun
16	24	23.3	22	12 090G62	090G62 human immun
17	24	23.3	22	12 090G60	090G60 human immun
18	24	23.3	22	12 090G66	090G66 human immun
19	24	23.3	22	12 090G62	090G62 human immun

20	24	23.3	22	12 090GFO	090GFO human immun
21	24	23.3	22	12 090G68	090G68 human immun
22	24	23.3	22	12 090G66	090G66 human immun
23	24	23.3	25	3 09P8S7	09P8S7 saccharomyc
24	24	23.3	15	10 093516	093516 arabidopsia
25	23	22.3	15	10 09S8W9	09S8W9 nicotiana a
26	23	22.3	20	1 09UWM8	09UWM8 halodactery
27	23	22.3	20	8 095947	095947 saccharomyc
28	23	22.3	23	8 092627	092627 dinodon sem
29	23	22.3	25	2 09X607	09X607 actinobacil
30	23	22.3	25	8 032966	032966 pinus thunb
31	22	21.4	10	12 090391	090391 tomato yell
32	22	21.4	11	9 038415	038415 bacterioph
33	22	21.4	12	2 052112	052112 actinobact
34	22	21.4	18	12 073605	073605 human immun
35	22	21.4	18	13 09PRS6	09PRS6 human immun
36	22	21.4	19	4 09UC21	09UC21 homo sapien
37	22	21.4	19	5 09TMR9	09TMR9 pyura stolo
38	22	21.4	19	8 092Y73	092Y73 microplitis
39	22	21.4	20	6 09TR77	09TR77 ovis aries
40	22	21.4	21	8 092YB7	092YB7 spinaria sp
41	22	21.4	21	8 09T2R4	09T2R4 solanum tub
42	22	21.4	22	3 09URC3	09URC3 phanerocha
43	22	21.4	22	3 09URC2	09URC2 phanerocha
44	22	21.4	22	4 09UCH8	09UCH8 homo sapien
45	22	21.4	24	2 045524	045524 bacillus su

ALIGNMENTS

RESULT 1	09KEL4	PRELIMINARY;	PRT;	14 AA.
ID	09KEL4			
AC	09KEL4			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	BH0838 PROTEIN.			
GN	BH0838.			
OS	Bacillus halodurans.			
OC	Bacteria: Firmicutes: Bacillus/Clostridium group;			
CC	Bacillus/staphylococcus group; Bacillus.			
OX	NCBI_TaxID=86665;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-C-125 / JCM 9153;			
RA	Takami H., Nakasone K., Takaki Y.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AP001510; BAB04557.1;			
SQ	SEQUENCE 14 AA: 1564 MW; C0359E12C13F5D4D CRC64;			
Query Match	26.28;	Score 27;	DB 2;	Length 14;
Best Local Similarity	40.08;	Pred. NO. 3.9e+02;		
Matches	4;	Conservative	4;	Mismatches 2;
				Indels 0;
				Gaps 0;
QY	5 VISNPRATHQ 14			
DB	5 VLNPATVIVE 14			
RESULT 2	09TWT4	PRELIMINARY;	PRT;	15 AA.
ID	09TWT4			
AC	09TWT4			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)			
DE	HEMUGLOBIN (FRAGMENT).			
OS	Lumbricus terrestris (Common earthworm).			
OC	Eukaryota: Metazoa: Annelida: Clitellata; Oligochaeta; Haplotaxidae;			
CC	Lumbricina; Lumbricidae; Lumbricus.			

OX NCBI_TaxID=6398;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93293879; PubMed=8514787;
 RA Oenby D.W., Zhu H., Schneider K., Beavis R.C., Chat B.T., Riggs A.F.;
 RT "The extracellular hemoglobin of the earthworm, Lumbricus terrestris.
 Determination of subunit stoichiometry."
 RL J. Biol. Chem. 268:13539-13547(1993).
 SQ SEQUENCE 15 AA; 1665 MW; F9B83FE0050D91F5 CRC64;

Query Match 26.28; Score 27; DB 5; Length 15;
 Best Local Similarity 50.08; Pred. No. 4.2e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 7 SNPAATHODI 16
 DB 2 SPRAQSHDEI 11

RESULT 3
 062645 PRELIMINARY; PRT; 17 AA.
 AC 062645;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DE MEMBRANE COFACTOR PROTEIN CD46 (FRAGMENT).
 OS Saguinus oedipus (Cotton-top tamarin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
 OX NCBI_TaxID=9490;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=98161817; PubMed=9494106;
 RA Murekani Y., Seta T., Kurita M., Fukui A., Ueda S., Nagasawa S.;
 RT Molecular cloning of membrane cofactor protein (MCP; CD46) on B95a
 cell, an Epstein-Barr virus-transformed marmoset B cell line: B95a-MCP
 is susceptible to infection by the CAM, but not the Nagahata strain of
 the measles virus."
 RL Biochem. J. 330:1351-1359(1998).
 DR EMBL; AB001991; BAA25631.1; -.
 KW Membrane.
 FT NON_TER 1 1
 SQ SEQUENCE 17 AA; 1983 MW; 1301A0ECD008F939 CRC64;

Query Match 25.28; Score 26; DB 6; Length 17;
 Best Local Similarity 42.98; Pred. No. 7.2e+02;
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 12 THODIDF 18
 DB 8 THREVN 14

RESULT 4
 046665 PRELIMINARY; PRT; 18 AA.
 ID 046665;
 AC 046665;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE GLUCOSE-6-PHOSPHATE DEHYDROGENASE (FRAGMENT).
 GN G6PD.
 OS Macropus robustus robustus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID=35350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97224585; PubMed=9060417;

RA Loebel D.A.F., Johnston P.G.;
 RT "Analysis of the intron-exon structure of the G6PD gene of the
 wallaroo (Macropus robustus) by polymerase chain reaction."
 RL Mamm. Genome 8:146-147(1997).
 DR EMBL; U53775; AAC48790.1; -.
 DR INTERPRO; IPR001282; -.
 DR PFAM; PF00479; G6PD; 1.
 FT NON_TER 1 1
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 1992 MW; CGD5981B528258FB CRC64;

Query Match 25.28; Score 26; DB 6; Length 18;
 Best Local Similarity 33.3%; Pred. No. 7.7e+02;
 Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 4 MVISNPAATHOD 15
 DB 1 VAMEKPASTNSD 12

RESULT 5
 09JIC5 PRELIMINARY; PRT; 19 AA.
 ID 09JIC5;
 AC 09JIC5;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE NAIP1 PROTEIN (FRAGMENT).
 GN NAIP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Gromey J.D., Dietrich W.F.;
 RT High Resolution Genetic and Physical Map of the Lgnl Interval in
 C57BL/6J Implicates Naip2 or Naip3 in Legionella pneumophila
 Pathogenesis."
 RL Genome Res. 0:0-0(2000).
 DR EMBL; AF240512; AAF81200.1; -.
 FT NON_TER 1 1
 FT NON_TER 19 19
 SQ SEQUENCE 19 AA; 1996 MW; 086C67103542478A CRC64;

Query Match 25.28; Score 26; DB 11; Length 19;
 Best Local Similarity 44.48; Pred. No. 8.2e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 8 NPAATHODI 16
 DB 10 DPAAVHSTV 18

RESULT 6
 084272 PRELIMINARY; PRT; 24 AA.
 ID 084272;
 AC 084272;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE E6 PROTEIN (FRAGMENT).
 GN Human papillomavirus type 19.
 OS Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
 OC NCBI_TaxID=10608;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88089511; PubMed=2826651;
 RA Krubke J., Kraus J., Dellus H., Chow L., Broker T.R., Ifner T.,
 Pfister H.;

RT "Genetic relationship among human papillomaviruses associated with
RT benign and malignant tumours of patients with epidermodysplasia
RT verruciformis."
RL J. Gen. Virol. 68:3091-3103(1987).
DR EMBL: D00204; BAA0143.1; -.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2562 MW; E0DD44BEA2811235 CRC64;

Query Match 25.2%; Score 26; DB 12; Length 24;
Best Local Similarity 28.6%; Pred. No. 1.1e+03;
Matches 4; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 6 ISNPAATHODIDFL 19
Db 1 MANNOATEEEIEIV 14

RESULT 7
ID Q10722 PRELIMINARY; PRT; 18 AA.
AC Q10722;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBlrel. 08, Last annotation update)
DE 30 KDA ANTIFUNGAL PROTEIN (FRAGMENT).
OS Engelmannia pinnatifida.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asterales;
OC euasterids II; Asterales; Asteraceae; Asteroideae; Heliantheae;
OC Engelmannia.
OX NCBI_TaxID=53580;
RN [1]
RP TISSUE=LEAF;
RC MEDLINE=96265034; PubMed=8670144;
RX Hyun Q.K., Borgmeyer J.R., Smith C.E., Bell L.D., Shah D.M.;
RA "Isolation and characterization of a 30 kDa protein with antifungal
RT activity from leaves of Engelmannia pinnatifida."
RL Biochem. J. 316:723-727(1996).
CC "- FUNCTION: EXHIBITS BROAD-SPECTRUM ANTIFUNGAL ACTIVITY."
KM Fungicide.
FT NON_TER 18
SQ SEQUENCE 18 AA; 2086 MW; A1F38FCFF0F451CA CRC64;

Query Match 24.3%; Score 25; DB 10; Length 18;
Best Local Similarity 40.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 FFRVVISNPA 10
Db 7 FETLALQXPA 16

RESULT 8
ID Q40183 PRELIMINARY; PRT; 19 AA.
AC Q40183;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)
DE NEGATIVELY LIGHT-REGULATED PROTEIN (Lg106) (FRAGMENT).
OS Lemna gibba (Swollen duckweed).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Araceae; Lemna.
OX NCBI_TaxID=4470;
RN [1]
RP SEQUENCE FROM N.A.
RA Okubara P.A., Flores S., Tobin E.M.;
RL Plant Mol. Biol. 11:673-681(1988).
DR EMBL: M35866; AAA33395.1; -.
FT NON_TER 1

SQ SEQUENCE 19 AA; 2153 MW; 975F6176AAB87387 CRC64;

Query Match 24.3%; Score 25; DB 10; Length 19;
Best Local Similarity 36.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 RMVISNPAATH 13
Db 1 RRIITOKASSH 11

RESULT 9
ID Q40181 PRELIMINARY; PRT; 21 AA.
AC Q40181;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)
DE SHORTEST ORF (FRAGMENT).
OS Lemna gibba (Swollen duckweed).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Araceae; Lemna.
OX NCBI_TaxID=4470;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L.;
RA Okubara P.A., Flores S., Tobin E.M.;
RL Plant Mol. Biol. 11:673-681(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=L.;
RA Okubara P.A., Flores S., Tobin E.M.;
RL Plant Mol. Biol. 15:955-956(1990).
DR EMBL: X14075; CAA3238.1; -.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2324 MW; B8C75F610576DB73 CRC64;

Query Match 24.3%; Score 25; DB 10; Length 21;
Best Local Similarity 36.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 RMVISNPAATH 13
Db 2 RRIITOKASSH 12

RESULT 10
ID Q62527 PRELIMINARY; PRT; 8 AA.
AC Q62527;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE TRANSTHYRETIN (PREALBUMIN) (FRAGMENT).
GN TTR.
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
OX NCBI_TaxID=10096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRET/ET;
RX MEDLINE=94319082; PubMed=8043949;
RA KO M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maezaki Y.,
Nadeau J.H.;
RT "Genetic mapping of 40 cDNA clones on the mouse genome by PCR."
RL Mamm. Genome 5:349-355(1994).
CC "- FUNCTION: THYROID HORMONE-BINDING PROTEIN. PROBABLY TRANSPORTS
CC THYROXINE FROM THE BLOODSTREAM TO THE BRAIN."
CC "- SUBUNIT: HOMOTETRAMER."
DR EMBL: U05689; AAB60461.1; -.

DR MGD; MGI:98865; Ttr.
 KW Albumin; Transport; Retinol-binding; Vitamin A; Thyroid hormone.
 FT NON_TER
 SQ SEQUENCE 8 AA: 828 MW: 9156C7645A2D2CD CRC64;

Query Match 23.3%; Score 24; DB 11; Length 8;
 Best Local Similarity 80.0%; Pred. No. 3.7e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 VISNP 9
 DB 2 VVSNP 6

RESULT 11

O9RA46 PRELIMINARY; PRT; 20 AA.
 AC O9RA46;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE 2-OXOGUTARATE:FERREDOXIN OXIDOREDUCTASE BETA SUBUNIT
 DE (EC 1.2.7.3) (FRAGMENT).
 OS Hydrogenobacter thermophilus.
 OC Bacteria; Aquificales; Aquificaceae; Hydrogenobacter.
 OX NCBI_Taxid=940;
 RN [1]
 FT NON_TER
 SQ SEQUENCE 20 AA: 2284 MW: BA040417DA7A3392 CRC64;

Query Match 23.3%; Score 24; DB 2; Length 20;
 Best Local Similarity 44.4%; Pred. No. 1.9e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 9 PATHODID 17
 DB 8 PADTKSDVE 16

RESULT 12

O69381 PRELIMINARY; PRT; 20 AA.
 AC O69381;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 08, Last annotation update)
 DE (HSV) DNA FOR IE MRNA-2 5' TERMINUS (FRAGMENT).
 OS Herpes simplex virus (type 2).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_Taxid=10310;
 RN [1]
 FT NON_TER
 SQ SEQUENCE 20 AA: 2180 MW: 11C39C0D07A447A9 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 20;

Best Local Similarity 83.3%; Pred. No. 1.9e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 DIDFLI 20
 DB 4 DIDMLI 9

RESULT 13

O9JG25 PRELIMINARY; PRT; 21 AA.
 AC O9JG25;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE ORF2 (FRAGMENT).
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_Taxid=68887;
 RN [1]
 FT NON_TER
 SQ SEQUENCE 21 AA: 1949 MW: 39AD537BFACABEC CRC64;

Query Match 23.3%; Score 24; DB 12; Length 21;
 Best Local Similarity 45.5%; Pred. No. 2e+03;
 Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 9 PATHODIDFL 19
 DB 5 PAGGEDADFL 15

RESULT 14

O68988 PRELIMINARY; PRT; 22 AA.
 AC O68988;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 08, Last annotation update)
 DE (HSV) DNA FOR IE MRNA-2 5' TERMINUS (FRAGMENT).
 OS human herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_Taxid=10298;
 RN [1]
 FT NON_TER
 SQ SEQUENCE 22 AA: 2411 MW: 03EBE1C22C0D07A4 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 22;
 Best Local Similarity 83.3%; Pred. No. 2.1e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 DIDFLI 20
 DB 4 DIDMLI 9

RESULT 15
ID 090G68 PRELIMINARY; PRT; 22 AA.
AC 090G68;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RP;
RX MEDLINE=99412391; PubMed=10482626;
RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
RA Bell J.E., Simmonds P.;
RT "Mosaic structure of the human immunodeficiency virus type 1 genome
RT infecting lymphoid cells and the brain: evidence for frequent in vivo
RT recombination events in the evolution of regional populations.";
RL J. Virol. 73:8720-8731(1999).
DR EMBL: AF174692; AAF00716.1; -.
DR INTERPRO: IPR000328; -.
DR PFAM: PF00517; GP41; 1.
KW Envelope protein.
FT NON_TER 1
SQ SEQUENCE 22 AA; 2704 MW; 86D008C65CAB43B4 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 22;
Best Local Similarity 37.5%; Pred. No. 2.1e+03;
Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 2 FRVYISNPATHDID 17
Db 3 FRAVHPRIRRGFE 18

RESULT 16
ID 090G62 PRELIMINARY; PRT; 22 AA.
AC 090G62;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RP;
RX MEDLINE=99412391; PubMed=10482626;
RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
RA Bell J.E., Simmonds P.;
RT "Mosaic structure of the human immunodeficiency virus type 1 genome
RT infecting lymphoid cells and the brain: evidence for frequent in vivo
RT recombination events in the evolution of regional populations.";
RL J. Virol. 73:8720-8731(1999).
DR EMBL: AF174692; AAF00723.1; -.
DR INTERPRO: IPR000328; -.
DR PFAM: PF00517; GP41; 1.
KW Envelope protein.
FT NON_TER 1
SQ SEQUENCE 22 AA; 2704 MW; 86D008C65CAB43B4 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 22;
Best Local Similarity 37.5%; Pred. No. 2.1e+03;

Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 2 FRVYISNPATHDID 17
Db 3 FRAVHPRIRRGFE 18

RESULT 17
ID 090G60 PRELIMINARY; PRT; 22 AA.
AC 090G60;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LF;
RX MEDLINE=99412391; PubMed=10482626;
RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
RA Bell J.E., Simmonds P.;
RT "Mosaic structure of the human immunodeficiency virus type 1 genome
RT infecting lymphoid cells and the brain: evidence for frequent in vivo
RT recombination events in the evolution of regional populations.";
RL J. Virol. 73:8720-8731(1999).
DR EMBL: AF174696; AAF00725.1; -.
DR INTERPRO: IPR000328; -.
DR PFAM: PF00517; GP41; 1.
KW Envelope protein.
FT NON_TER 1
SQ SEQUENCE 22 AA; 2704 MW; 86D008C65CAB43B4 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 22;
Best Local Similarity 37.5%; Pred. No. 2.1e+03;
Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 2 FRVYISNPATHDID 17
Db 3 FRAVHPRIRRGFE 18

RESULT 18
ID 090G66 PRELIMINARY; PRT; 22 AA.
AC 090G66;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LF;
RX MEDLINE=99412391; PubMed=10482626;
RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
RA Bell J.E., Simmonds P.;
RT "Mosaic structure of the human immunodeficiency virus type 1 genome
RT infecting lymphoid cells and the brain: evidence for frequent in vivo
RT recombination events in the evolution of regional populations.";
RL J. Virol. 73:8720-8731(1999).
DR EMBL: AF174698; AAF00729.1; -.
DR INTERPRO: IPR000328; -.
DR PFAM: PF00517; GP41; 1.
KW Envelope protein.
FT NON_TER 1

SQ SEQUENCE 22 AA: 2704 MW: 86D008C65CAB43B4 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 22;

Best Local Similarity 37.5%; Pred. No. 2.1e+03; Indels 0; Gaps 0;

OY 2 FRVVISNPAATHODID 17
DB 3 FRVVIHPRIRIOGFE 18

RESULT 19

OQOGF2 PRELIMINARY; PRT; 22 AA.

AC OQOGF2: 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_Taxid-11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-LN1;

RA MEDLINE-99412391; PubMed-10482626;

RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,

RA Bell J.E., Simmonds P.;

RT "Mosaic structure of the human immunodeficiency virus type 1 genome

infected lymphoid cells and the brain: evidence for frequent in vivo

recombination events in the evolution of regional populations.";

RL J. Virol. 73:8720-8731(1999).

DR EMBL: AF174700; AAF00733.1;

DR INTERPRO: IPR000328;

DR PFAM: PF00517; GP41; 1.

KW Envelope protein.

FT NON_TER 1

FT SEQUENCE 22 AA: 2676 MW: 86D008C6441C13B4 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 22;

Best Local Similarity 37.5%; Pred. No. 2.1e+03;

Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 2 FRVVISNPAATHODID 17

DB 3 FRVVIHPRIRIOGFE 18

RESULT 20

OQOGF0 PRELIMINARY; PRT; 22 AA.

AC OQOGF0: 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_Taxid-11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-LN2;

RA MEDLINE-99412391; PubMed-10482626;

RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,

RA Bell J.E., Simmonds P.;

RT "Mosaic structure of the human immunodeficiency virus type 1 genome

infected lymphoid cells and the brain: evidence for frequent in vivo

recombination events in the evolution of regional populations.";

RL J. Virol. 73:8720-8731(1999).

DR EMBL: AF174701; AAF00735.1; -.

DR INTERPRO: IPR000328; -.

DR PFAM: PF00517; GP41; 1.

DR Envelope protein.

FT NON_TER 1

FT SEQUENCE 22 AA: 2676 MW: 86D008C6441C13B4 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 22;

Best Local Similarity 37.5%; Pred. No. 2.1e+03;

Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 2 FRVVISNPAATHODID 17

DB 3 FRVVIHPRIRIOGFE 18

RESULT 21

OQOGF8 PRELIMINARY; PRT; 22 AA.

AC OQOGF8: 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_Taxid-11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Cp-1;

RA MEDLINE-99412391; PubMed-10482626;

RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,

RA Bell J.E., Simmonds P.;

RT "Mosaic structure of the human immunodeficiency virus type 1 genome

infected lymphoid cells and the brain: evidence for frequent in vivo

recombination events in the evolution of regional populations.";

RL J. Virol. 73:8720-8731(1999).

DR EMBL: AF174702; AAF00737.1; -.

DR INTERPRO: IPR000328; -.

DR PFAM: PF00517; GP41; 1.

KW Envelope protein.

FT NON_TER 1

FT SEQUENCE 22 AA: 2676 MW: 86D008C6441C13B4 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 22;

Best Local Similarity 37.5%; Pred. No. 2.1e+03;

Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 2 FRVVISNPAATHODID 17

DB 3 FRVVIHPRIRIOGFE 18

RESULT 22

OQOGF6 PRELIMINARY; PRT; 22 AA.

AC OQOGF6: 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_Taxid-11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Cp-2;

RA MEDLINE-99412391; PubMed-10482626;

RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,

RA Bell J.E., Simmonds P.;
 RT "Mosaic structure of the human immunodeficiency virus type 1 genome
 RT infecting lymphoid cells and the brain: evidence for frequent in vivo
 RT recombination events in the evolution of regional populations.";
 RL J. Virol. 73:8720-8731(1999).
 DR EMBL: AF174703; AAF00739.1;
 DR INTERPRO: IPR000328;
 DR PFM: PFO0517; GP41; 1.
 KW Envelope protein.
 FT NON_TER 1
 SQ SEQUENCE 22 AA; 2662 MW; 86D008C640C733B4 CRC64;

Query Match 23.3%; Score 24; DB 12; Length 22;
 Best Local Similarity 37.5%; Pred. No. 2.1e+03;
 Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 FRWVSNPATHQDID 17
 DB 3 FRAVHIFRIRIOCFE 18

RESULT 23
 Q9P8S7 PRELIMINARY; PRT; 25 AA.
 AC Q9P8S7;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE ALCOHOL DEHYDROGENASE ISOLYME 2 (FRAGMENT).
 OS Saccharomyces douglasi (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=46617;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Young E.T., Sloan J., Miller B., Li N., van Riper K., Dombeck K.M.;
 RT "Evolution of a glucose-regulated ADH gene in the genus
 RT Saccharomyces";
 RL Gene 245:299-309(2000).
 DR EMBL: AF218307; AAF43643.1;
 FT NON_TER 25
 SQ SEQUENCE 25 AA; 2860 MW; 219627E5620191F9 CRC64;

Query Match 23.3%; Score 24; DB 3; Length 25;
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 7 SNPAATHQDI 16
 DB 15 SNGKLEHKDI 24

RESULT 24
 P93516 PRELIMINARY; PRT; 15 AA.
 AC P93516;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DE COPPER AMINE OXIDASE (FRAGMENT).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurossids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Woller S.G., McPherson M.J.;
 RL Biochem. Soc. Trans. 23:0-0(1995).
 DR EMBL: S82298; AAB37691.1;
 DR INTERPRO: IPR000269;

DR PFM: PFO1179; Cu_amine_oxid; 1.
 FT NON_TER 1
 SQ SEQUENCE 15 AA; 1765 MW; 1A836E041E00C5CE CRC64;

Query Match 22.3%; Score 23; DB 10; Length 15;
 Best Local Similarity 45.5%; Pred. No. 2.1e+03;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 AATHQDIDFLI 20
 DB 3 AATVGNVDYII 13

RESULT 25
 Q9S8W9 PRELIMINARY; PRT; 15 AA.
 AC Q9S8W9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE SELF-INCOMPATIBILITY S3 GLYCOPROTEIN (FRAGMENT).
 OS Nicotiana glauca (Winged tobacco) (Persian tobacco).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4087;
 RN [1]
 RP SEQUENCE.
 RA MEDLINP-92404717; PubMed=2535548;
 RA Jahn W., Batterman M.P., Clarke A.E., Moritz R.L., Simpson R.J.;
 RT "Identification, isolation, and N-terminal sequencing of style
 RT glycoproteins associated with self-incompatibility in Nicotiana
 RT glauca";
 RL Plant Cell 1:493-499(1989).
 SQ SEQUENCE 15 AA; 1814 MW; D9F5D7B05B215937 CRC64;

Query Match 22.3%; Score 23; DB 10; Length 15;
 Best Local Similarity 36.4%; Pred. No. 2.1e+03;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 FFRWVSNPAA 11
 DB 4 YMOVLQWPPA 14

Search completed: February 5, 2001, 10:55:30
 Job time: 919 sec

Tue Feb 6 08:48:27 2001

us-08-981-824-7.rspt

RN [1]
 RP SEQUENCE.
 RX MEDLINE-91369922: PubMed-1892823;
 RA Zhang J.-H., Kurtz D.M. Jr.:
 RT "Two distinct subunits of hemerythrin from the brachiopod Lingula
 RL reveal: an apparent requirement for cooperatively in O2 binding."
 CC Biochemistry 30:9121-9124(1991).
 CC -1- FUNCTION: HEMERYTHRIN IS A RESPIRATORY PROTEIN IN BLOOD CELLS OF
 CC CERTAIN MARINE WORMS. THE OXYGEN-BINDING SITE IN EACH CHAIN
 CC CONTAINS TWO IRON ATOMS.
 CC -1- SUBUNIT: OCTAMER COMPOSED OF TWO TYPES OF CHAINS: ALPHA AND BETA.
 CC -1- SIMILARITY: BELONGS TO THE HEMERYTHRIN FAMILY.
 DR INTERPRO: IPR002063; -
 DR PROSITE: PS00550; HEMERYTHRINS; PARTIAL.
 KW Oxygen transport; Metal-binding; Iron.
 FT METAL 24 24 IRON 1 (BY SIMILARITY).
 FT NON_TER 24 24
 SO SEQUENCE 24 AA: 2787 MW: 2F0121054632028A CRC64;

Query Match 24.3%; Score 25; DB 1; Length 24;
 Best Local Similarity 45.5%; Pred. No. 3.5e+02;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 SNPATTMODID 17
 DB 11 SDFATTYENID 21

RESULT 3
 PK12_SQUTU STANDARD: PRT: 25 AA.
 ID PK12_SQUTU
 AC P24744;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-MAR-1992 (Rel. 21, Last annotation update)
 DE KUNITZ-TYPE INHIBITOR-2 (PK1-2) (FRAGMENT).
 OS Solanum tuberosum (Potato).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Asteridae I;
 CC Solanales; Solanaceae; Solanum.
 RN [1]
 RP SEQUENCE.
 RC STRAIN-CV. RUSSET BURBANK; TISSUE-TUBER;
 RA Walsh T.A., Twilchell W.P.;
 RT "Two Kunitz-type proteinase inhibitors from potato tubers."
 RL Plant Physiol. 97:15-18(1991).
 CC -1- FUNCTION: POTENT INHIBITOR OF SUBTILISIN. MODERATE INHIBITOR
 CC OF TRYPSIN AND CHYMOTRYPSIN.
 CC -1- TISSUE SPECIFICITY: CONTEXT OF POTATO TUBER.
 CC -1- SIMILARITY: TO SOYBEAN TRYPSIN INHIBITOR (KUNITZ) FAMILY OF
 CC PROTEASE INHIBITOR.
 DR INTERPRO: IPR002160; -
 DR PFAM: PF00197; Kunitz legume; 1.
 DR PROSITE: PS00283; SOYBEAN_KUNITZ; 1.
 KW Serine protease inhibitor.
 FT NON_TER 25 25
 SO SEQUENCE 25 AA: 2920 MW: 1F5607405921DFDC CRC64;

Query Match 24.3%; Score 25; DB 1; Length 25;
 Best Local Similarity 57.1%; Pred. No. 3.7e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 RMVSNP 9
 DB 19 RYINNP 25

RESULT 4
 POQA_PSEFL STANDARD: PRT: 24 AA.
 ID POQA_PSEFL
 AC P55171;

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE COENZYME PQO SYNTHESIS PROTEIN A.
 GN POQA.
 OS Pseudomonas fluorescens.
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CHAO;
 RX MEDLINE-96064397; PubMed-8526497;
 RA Schneider U., Keel C., Defago G., Haas D.;
 RT "Tis-directed cloning of pqr genes from Pseudomonas fluorescens CHAO:
 RT mutational inactivation of the genes results in overproduction of the
 RT antibiotic pyoluteorin."
 RL Appl. Environ. Microbiol. 61:3856-3864(1995).
 CC -1- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)
 CC BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ
 CC SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH
 CC ARE NECESSARY FOR THE SYNTHESIS.
 CC -1- SIMILARITY: TO OTHER BACTERIAL PQQA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X87299; CAA60732.1; -
 DR EMBL: X87299; CAA60731.1; ALT_INIT.
 DR PQQ.
 KW PQO.
 SO SEQUENCE 24 AA: 2871 MW: ED13A879B2EA3E8E CRC64;

Query Match 23.3%; Score 24; DB 1; Length 24;
 Best Local Similarity 46.7%; Pred. No. 5.2e+02;
 Matches 7; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 WVISNPATMODIDP 18
 DB 1 MTWSKPATYDRIGF 15

RESULT 5
 BURK_CLOPA STANDARD: PRT: 9 AA.
 ID BURK_CLOPA
 AC P81337;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE BUTYRATE KINASE (EC 2.7.2.7) (BK) (CP 38) (FRAGMENT).
 GN BURK.
 OS Clostridium pasteurianum.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 CC Clostridium.
 RN [1]
 RP SEQUENCE.
 RC STRAIN-W5.
 RX MEDLINE-96291870; PubMed-9629918;
 RA Flengard R., Skjeldal L.;
 RT "Two-dimensional gel electrophoresis separation and N-terminal
 RT sequence analysis of proteins from Clostridium pasteurianum W5."
 RL Electrophoresis 19:802-806(1998).
 CC -1- FUNCTION: CATALYZES THE CONVERSION OF BUTYRYL-COA THROUGH BUTYRYL
 CC PHOSPHATE TO BUTYRATE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + 2-BUTANOATE -> ADP + BUTANOYL PHOSPHATE.
 CC -1- PATHWAY: INVOLVED IN THE ACIDOGENIC PHASE OF FERMENTATION.
 CC -1- SIMILARITY: BELONGS TO THE ACETOKINASE FAMILY.
 DR INTERPRO: IPR000890; -
 DR PROSITE: PS01075; ACETATE_KINASE_1; PARTIAL.


```

DR PROSITE; PS01076; ACETATE_KINASE_2; PARTIAL.
KW TRANSFERASE; KINASE.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1104 MW; 055D40505727232B CRC64;

Query Match
Best Local Similarity 22.3%; Score 23; DB 1; Length 9;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 FRWISNP 9
DB 2 YKLIINP 9

RESULT 6
UC29_MAIZE STANDARD; PRT; 15 AA.
AC P80635;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLOGICAL COLEOPTILE (SPOT 45)
DE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN [1]
RP TISSUE-COLEOPTILE;
RA Touzet P., Niccardi F., Morin C., Damerwal C., Huet J.-C.,
RA Perinelli J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 4.9, ITS MW IS: 37.6 KDA.
DR MAIZE-2DPAGE; P80635; COLEOPTILE.
DR MAIZEDB; 123960; -.
FT NON_TER 15
SQ SEQUENCE 15 AA; 1679 MW; 3D53086B16018BC1 CRC64;

Query Match
Best Local Similarity 22.3%; Score 23; DB 1; Length 15;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 DIDFL 19
DB 11 DIDYL 15

RESULT 7
SRD_ATREN STANDARD; PRT; 21 AA.
AC P13211;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE SARAFOTOXIN-D (S6D) (SRFX-D).
OS Atractaspis engaddensis (Israeli burrowing asp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosteiformes; Serranidae; Scleroglossa; Serpentes; Colubroidae;
OC Atractaspididae; Atractaspis.
RN [1]
RP SEQUENCE.
RC TISSUE-VENOM;
RX MEDLINE-90033283; Pubmed-2509240;
RA Bohl A., Wollberg Z., Flemlinger G.;
RT "SRFX-d, a new native peptide of the endothelin/sarafotoxin family.";
RL FEBS Lett. 256:1-3(1989)
CC -1- FUNCTION: VASOCONSTRICTOR ACTIVITY. THESE TOXINS CAUSE CARDIAC

```

```

CC ARREST PROBABLY AS A RESULT OF CORONARY VASOPLASM.
CC -1- MISCELLANEOUS: LD(50) IS 0.35 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE ENDOTHELIN/SARAFOTOXIN FAMILY.
KW PIR; S06145; S06145.
DR HSSP; P13208; 1SRB.
DR INTERPRO; IPR001928; -.
DR PRAM; PF00332; endothelin; 1.
DR PRINTS; PR00365; ENDOTHELIN.
DR PROSITE; PS00270; ENDOTHELIN; 1.
KW Venom; Vasoconstrictor; Toxin.
FT DISULFID 1 15
FT DISULFID 3 11 BY SIMILARITY.
SQ SEQUENCE 21 AA; 2596 MW; 83A8A04D1D536AE2 CRC64;

Query Match
Best Local Similarity 22.3%; Score 23; DB 1; Length 21;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 HODI 16
DB 16 HODI 19

RESULT 8
ARGD_SALTY STANDARD; PRT; 17 AA.
ID ARGD_SALTY
AC P40732;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE ACETYLORNITHINE AMINOTRANSFERASE (EC 2.6.1.11) (ACOAT) (FRAGMENT).
GN ARGD OR DTU.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90094246; Pubmed-2403545;
RA Tran P.V., Bannor T.A., Doktor S.Z., Nichols B.P.;
RT "Chromosomal organization and expression of Escherichia coli paba.";
RL J. Bacteriol. 172:397-410(1990).
CC -1- CATALYTIC ACTIVITY: N2-ACETYL-L-ORNITHINE + 2-OXOGLOUTAMATE -
CC N-ACETYL-L-GLUTAMATE 5-SEMIALDEHYDE + L-GLUTAMATE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: FOURTH STEP IN ARGinine BIOSYNTHESIS.
CC -1- MISCELLANEOUS: THE REACTION CATALYZED BY ACOAT IS HIGHLY
CC REVERSIBLE. MOREOVER THIS ENZYME MAY TRANSMUTATE ORNITHINE.
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32354; AAA24265.1; -.
DR EMBL; M32355; AAA27178.1; -.
DR STYGENE; SG10488; ARGD.
DR INTERPRO; IPR000954; -.
DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
KW Transferase; Aminotransferase; Pyridoxal phosphate;
KW Arginine biosynthesis.
FT NON_TER 17
SQ SEQUENCE 17 AA; 1909 MW; 03DD865FE9D16B9 CRC64;

Query Match
Best Local Similarity 21.4%; Score 22; DB 1; Length 17;
Matches 6; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

```

QY 4 VVISNPAATHQDIDFLI 20
 DB 1 MAIEQFAIRATFEVY 17

RESULT 9

RL23_HALCU STANDARD: PRT: 18 AA.
 AC P05975;
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L23P (HL31) (FRAGMENT).
 GN RPL23P.
 OS Halobacterium cutribdum.
 OC Archaea: Euryarchaeota; Halobacteriales; Halobacteriaceae;
 CC Halobacterium.
 RN [1]
 RP SEQUENCE.

RA MEDLINE-84282108; PubMed-6467081;
 RA Matheson A.T., Yaguchi M., Christensen P., Rollin C.F., Hasnain S.;
 RT "Purification, properties, and N-terminal amino acid sequence of
 RT certain 50S ribosomal subunit proteins from the archaeobacterium
 RT Halobacterium cutribdum.";
 RL Can. J. Biochem. Cell Biol. 62:426-433(1984).
 CC -1- FUNCTION: BINDS TO A SPECIFIC REGION ON THE 23S RNA
 (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
 DR PIR. S08559; S08559.
 DR INTERPRO: IPR001014;
 DR PROSITE: PS00050; RIBOSOMAL_L23; PARTIAL.
 DR KW Ribosomal protein; rRNA-binding.
 FT NON_TER 18
 FT SEQUENCE 18 AA; 2041 MW; 361DC53F8B4D4C CRC64;

Query Match 21.4%; Score 22; DB 1; Length 18;
 Best Local Similarity 30.8%; Pred. No. 8.4e+02;
 Matches 4; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 VVISNPAATHQDIDFLI 17
 DB 3 IIDYPLVTEKAMD 15

RESULT 10
 FIBA_CEREL STANDARD: PRT: 19 AA.
 AC P14446;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-JAN-1990 (Rel. 13, Last annotation update)
 DE FIBRINOPEPTIDE A.
 OS Cerus elaphus (red deer), and Cerus elaphus nelsoni (American elk).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
 CC Cervidae; Cervinae; Cervus.
 RN [1]
 RP SEQUENCE.

RC SPECIES=C. ELAPHUS;
 RA Blomback B., Blomback M., Grondahl N.J.;
 RT "Studies on fibrinopeptides from mammals.";
 RL Acta Chem. Scand. 19:1789-1791(1965).
 RN [2]
 RP SEQUENCE.

RC SPECIES=C. E. NELSONI;
 RA Mross G.A., Doolittle R.F.;
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
 RL Arch. Biochem. Biophys. 122:674-684(1967).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.

CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

DR HSP; P02671; IBBR.
 KW Blood coagulation; Plasma.
 FT NON_TER 19
 FT SEQUENCE 19 AA; 1808 MW; 9BA54C26873B59C5 CRC64;

Query Match 21.4%; Score 22; DB 1; Length 19;
 Best Local Similarity 46.2%; Pred. No. 8.9e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

QY 7 SNPAATHQDIDFLI 19
 DB 4 SDPASS----DFL 12

RESULT 11

MIFH_TRISP STANDARD: PRT: 19 AA.
 AC P81529;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE L-DOPACHROME-METHYL ESTER TAUTOMERASE (MACROPHAGE MIGRATION INHIBITORY
 DE FACTOR HOMOLOG) (FRAGMENT).
 OS Trichinella spiralis.
 OS Eukaryota; Metazoa; Nematoda; Enoplia; Trichocephalida;
 CC Trichocephalina; Trichinelloidea; Trichinellidae; Trichinella.
 RN [1]
 RP SEQUENCE.

RA Pennock U.L., Behnke J.M., Bickle O.D., Devaney E., Grens R.K.,
 RA Isaac R.E., Joshua G.W.P., Selkirk M.E., Zhang Y., Meyer D.J.;
 RT "Rapid purification and characterization of L-dopamine-methyl-ester
 RT tautomerase (macrophage migration inhibitory factor) from Trichinella
 RT spiralis, Trichinella muris and Brugia pahangi.";
 RL Biochem. J. 335:495-498(1998).
 CC -1- FUNCTION: TAUTOMERIZATION OF THE METHYL ESTER OF L-DOPACHROME.
 CC -1- INDUCTION: INHIBITED BY FREE FATTY ACIDS AND HAEMATIN.
 CC -1- SIMILARITY: BELONGS TO THE MIF FAMILY.
 DR INTERPRO: IPR001398;
 DR PROSITE: PS01158; MIF; PARTIAL.
 DR CYCLOPINE.
 FT NON_TER 19
 FT SEQUENCE 19 AA; 2106 MW; F7CA05F112A628D CRC64;

Query Match 21.4%; Score 22; DB 1; Length 19;
 Best Local Similarity 46.2%; Pred. No. 8.9e+02;
 Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 7 SNPAATHQDIDFLI 19
 DB 7 TIRKATDVPSDFL 19

RESULT 12
 BTX_ATRBI STANDARD: PRT: 21 AA.
 AC P80163;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE BIROROXIN (BTX).
 OS Attractaspis bibroni (Southern bibron's) (Stiletto snake).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 CC Attractaspididae; Attractaspis.

```

RN [1]
RP SEQUENCE.
RC TISSUE-VENOM;
RX MEDLINE-93106214; PubMed-8416802;
RA Becker A., Dowdle E.B., Hechler U., Kauser K., Donner P.,
  Schlenker W.-D.;
RT "Bibrotaxin, a novel member of the endothelin/sarafotoxin peptide
  family, from the venom of the burrowing asp Atractaspis bilineata.";
RL FEBS Lett. 315:100-103(1993).
CC -1- FUNCTION: VASOCONSTRICTOR ACTIVITY. THESE TOXINS CAUSE CARDIAC
  ARREST PROBABLY AS A RESULT OF CORONARY VASOPASM.
CC -1- SIMILARITY: BELONGS TO THE ENDOTHELIN/SARAFOTOXIN FAMILY.
DR HSP: S27039; S27039.
DR HSP: P13208; ISRB.
DR INTERPRO: IPR001928;
DR PFAM: PF00322; endothelin; 1.
DR PRINTS: PR00365; ENDOTHELIN.
DR PROSITE: PS00270; ENDOTHELIN; 1.
KM Venom. Vasoconstrictor; Toxin.
FT DISULFID 1 15 BY SIMILARITY.
FT DISULFID 3 11 BY SIMILARITY.
SQ SEQUENCE 21 AA; 2511 MW; 83A5DF81D036AE2 CRC64;

Query Match 21.48; Score 22; DB 1; Length 21;
Best Local Similarity 75.08; Pred. No. 9.9e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 HQDI 16
  111;
Db 16 HQDV 19

RESULT 13
AROE_BUCAI STANDARD; PRT; 21 AA.
AC 044607;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SHIKIMATE 5-DEHYDROGENASE (EC 1.1.1.25) (FRAGMENT).
AROE OR BU493.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
  symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95212914; PubMed-7535281;
RA Roubbaksh D., Baumann P.;
RT "Characterization of a putative 23S-5S rRNA operon of Buchnera
  aphidicola (endosymbiont of aphids) unlinked to the 16S rRNA-encoding
  gene.";
RL Gene 155:107-112(1995).
CC -1- CATALYTIC ACTIVITY: SHIKIMATE + NADP(+) -> 5-DEHYDROSHIKIMATE +
  NADPH.
CC -1- PATHWAY: FOURTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE
  BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SIMILARITY: BELONGS TO THE SHIKIMATE DEHYDROGENASE FAMILY.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
  FRAMESHIFT IN POSITION 13.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U10496; AAA79125.1; ALT_FRAME.
KW Aromatic amino acid biosynthesis; Oxidoreductase; NADP.
FT NON_TER 1 1
SQ SEQUENCE 21 AA; 2590 MW; 73AE77462AB812E6 CRC64;

```

```

Query Match 20.48; Score 21; DB 1; Length 21;
Best Local Similarity 50.08; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 15 DIDFLI 20
  111;
Db 11 EIDYII 16

RESULT 14
ECD1_LYMDI STANDARD; PRT; 21 AA.
ID ECD1_LYMDI
AC P80936; P55898;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TESTIS ECDYSTOTROPIN PEPTIDE 1 (TE).
OS Lymantria dispar (Gypsy moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
  Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
  OC Noctuoidea; Lymantriidae; Lymantria.
RN [1]
RP SEQUENCE.
RC TISSUE-BRAIN;
RA Wagner R.M., Loeb M.J., Kochansky J.P., Gelman D.B., Lusby W.R.,
  Bell R.A.;
RT "Identification and characterization of an ecdysiotropic peptide from
  brain extracts of the gypsy moth, Lymantria dispar.";
RL Arch. Insect Biochem. Physiol. 34:175-189(1997).
CC -1- FUNCTION: START OR BOOST ECDYSTEROID SYNTHESIS IN TESTIS OF LARVAE
  AND PUPAE.
SQ SEQUENCE 21 AA; 2474 MW; E6E819624D1DEB62 CRC64;

Query Match 20.48; Score 21; DB 1; Length 21;
Best Local Similarity 36.48; Pred. No. 1.5e+03;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 NPAATHODIDF 18
  111;
Db 11 NDAUNNEVLD 21

RESULT 15
IF2G_PIG STANDARD; PRT; 22 AA.
ID IF2G_PIG
AC P20461;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE EUKARYOTIC TRANSLATION INITIATION FACTOR 2 GAMMA SUBUNIT (EIF-2-
  DE GAMMA) (FRAGMENT).
CN EIF2G3 OR EIF2G.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
RN [1]
RP SEQUENCE.
RC TISSUE-LIVER.
RA Suzuki H., Mukoyama E.B.;
RT "Pig liver translational initiation factor eif-2: N-terminal amino
  acid sequences of alpha and gamma subunits and the phosphorylation
  site structure.";
RL Agric. Biol. Chem. 52:1397-1408(1988).
CC -1- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS
  BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR tRNA. THIS
  COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY RNA BINDING
  TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL
  SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY
  HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP
  BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER

```

```

CC ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP
CC BY WAY OF A REACTION CATALYZED BY EIF-2B.
CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA, AND A GAMMA
CC CHAIN.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EIF2G SUBFAMILY.
CC PIR: P70052; P70052.
CC Initiation factor; Protein biosynthesis; GTP-binding.
CC NON_TER 22
CC SEQUENCE 22 AA; 2203 MM; 428BA7D77D18B03C CRC64;
SQ
Query Match 20.4%; Score 21; DB 1; Length 22;
Best Local Similarity 25.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 4 MVISNPAATHOD 15
DB 7 VTIGDPSLXKMD 18

RESULT 16
MOTI_CANFA STANDARD; PRT; 22 AA.
AC P19863:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE MOTILIN.
GN MLN.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE.
RC TISSUE-INTERESTINE; PubMed-6844663;
RX MEDLINE-83195948;
RA Poltras P., Reeve J.R. Jr., Hunkapiller M.W., Hood L.E., Walsh J.H.;
RT "Purification and characterization of canine intestinal motilin.";
RL Regul. Pept. 5:197-208(1983).
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE REGULATION OF
CC INTERDIGESTIVE GASTROINTESTINAL MOTILITY AND INDIRECTLY CAUSES
CC RHYTHMIC CONTRACTION OF DUODENAL AND COLONIC SMOOTH MUSCLE.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE MOTILIN FAMILY.
CC PIR: S00189; S00189.
CC DR PIR; A60313; A60313.
CC KW Hormone.
CC FT UNSURE
CC SEQUENCE 1 1
SQ SEQUENCE 22 AA; 2685 MM; 4BECB840ABE0639F CRC64;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87215944; PubMed-3555844;
RA Werner M., Feller A., Messenguy F., Pierard A.;
RT "The leader peptide of yeast gene CPA1 is essential for the
RT translational repression of its expression.";
RL Cell 49:805-813(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-85101411; PubMed-3681260;
RA Werner M., Feller A., Pierard A.;
RT "Nucleotide sequence of yeast gene CP A1 encoding the small subunit
RT of arginine-pyruvate carboxyl-phosphate synthetase. Homology of the
RT deduced amino acid sequence to other glutamine amidotransferases.";
RL Eur. J. Biochem. 146:371-381(1985).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE-97298310; PubMed-9153758;
RA Potrey R., Cziepluch C., Tobiasch E., Pujol A., Kordes E.,
RA Jauniaux J.-C.;
RT "Sequence and analysis of a 36.2 kb fragment from the right arm of
RT yeast chromosome XV reveals 19 open reading frames including SNF2 (5'
RT end), CPA1, SLX41, a putative transport ATPase, a putative ribosomal
RT protein and an SNF2 homologue.";
RL Yeast 13:479-482(1997).
CC -1- SIMILARITY: TO A HYPOTHETICAL PROTEIN IN ARG-2 5' REGION FROM
CC NEUROSPORA CRASSA.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC CC
CC EMBL; M16590; AAA34526.1; -
CC DR EMBL; X01764; CAN25904.1; -
CC DR EMBL; 275210; CAN99620.1; -
CC DR PIR; A26965; A26965.
CC DR PIR; S25434; S25434.
CC SCD; S0005828; YOR302W.
CC KW Leader Peptide.
CC SEQUENCE 25 AA; 2979 MM; 01D494C8AFEA4AA CRC64;
SQ
Query Match 20.4%; Score 21; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 6 ISNPAATHOD 15
DB 4 LNSOYTCOD 13

RESULT 18
SMBP_RAT STANDARD; PRT; 25 AA.
ID SMBP_RAT
AC P80968;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SM-11044 BINDING PROTEIN (FRAGMENTS).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE.
RC STRAIN-WISTAR.
RX MEDLINE-97407910; PubMed-9261134;
RA Sugisawa T., Matsuzaki-Fujita M., Guillaume J.-L., Camoin L.,
RA Morooka S., Strosberg A.D.;

```

RT "Characterization of a novel Iodocyanopindolol and SM-11044 binding
RT protein, which may mediate relaxation of depolarized rat colon
RT tonsus";
RL J. Biol. Chem. 272:21244-21252(1997).
CC -1- FUNCTION: MAY MEDIATE RELAXATION OF DEPOLARIZED COLON TONSUS.
CC IT BINDS IODOCYANOPINDOLOL AND SM-11044.
CC -1- MISCELLANEOUS: THE ORDER OF THE FRAGMENTS IS UNKNOWN.
FT NON_TER 1 1
FT UNSURE 6 6 OR Y.
FT NON_CONS 18 19
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3177 MW; D14FCB9B78C2CB CRC64;

Query Match 20.4%; Score 21; DB 1; Length 25;
Best Local Similarity 28.6%; Pred. No. 1.8e+03;
Matches 4; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 FFRVISPATRHQ 14
DB 2 FVFFILADPARFQ 15

RESULT 19
LCMS_LEDMA STANDARD; PRT; 10 AA.
AC P21144; P41497;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE LEUCOMYOSUPPRESSIN (LMS) (LEM-MS).
OC Leucophaea maderae (Madelira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoides; Blaberidae; Leucophaea.
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE-HEAD;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT Isolation, primary structure and synthesis of leucomyosuppressin,
an insect neuropeptide that inhibits spontaneous contractions of the
RT cockroach hindgut";
RL Comp. Biochem. Physiol. 85C:329-333(1986).
CC -1- FUNCTION: INHIBITS THE SPONTANEOUS CONTRACTIONS OF COCKROACH
CC PROCTODUM (HINDGUT).
KM Neuropeptide; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1275 MW; D3C45229D2C1EAB2 CRC64;

Query Match 19.4%; Score 20; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 9.5e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 QDID 17
DB 1 QDVD 4

RESULT 20
ALUS_MANSE STANDARD; PRT; 15 AA.
ID ALUS_MANSE
AC P42559;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ALLATOSTATIN (MAS-AS).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Dictyyla;
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
RN [1]

RP SEQUENCE.
RC TISSUE-HEAD;
RX MEDLINE-92052112; PubMed-1946359;
RA Kramer S.J., Toschi A., Miller C.A., Kataoka H., Quistad G.B.,
RL J.P., Carney R.L., Schooley D.A.;
RT Identification of an allatostatin from the tobacco hornworm Manduca
RT sexta";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9458-9462(1991).
CC -1- FUNCTION: STRONGLY INHIBITS JUVENILE HORMONE BIOSYNTHESIS IN VITRO
CC BY THE CORPORA ALLATA FROM FIFTH-STADIUM LARVAE AND ADULT FEMALES.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KM Neuropeptide.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 15 AA; 1908 MW; 1605B77CDEBC838E CRC64;

Query Match 19.4%; Score 20; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 FFRVISP 9
DB 4 FROCFNP 11

RESULT 21
UP37_UPEMJ STANDARD; PRT; 17 AA.
ID UP37_UPEMJ
AC P82044;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE UPERIN 3.7.
OS Uperoleia mjobergii (Australian toadlet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Myobatrachidae;
OC Uperoleia.
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-SKIN SECRETION;
RA Bradford A.M., Bowie J.H., Tyler M.J., Wallace J.C.;
RT New antibiotic uperolein peptides from the dorsal glands of the
RT Australian toadlet Uperoleia mjobergii";
RL Aust. J. Chem. 49:1325-1331(1996).
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW-1844; METHOD-FAB.
KM Amphibian skin; Amidation.
FT MOD_RES 1 1 AMIDATION.
SQ SEQUENCE 17 AA; 1845 MW; 655B748FB438B67B CRC64;

Query Match 19.4%; Score 20; DB 1; Length 17;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 FFRVISP 7
DB 6 FFRIVS 11

RESULT 22
PSAE_CUCSA STANDARD; PRT; 19 AA.
ID PSAE_CUCSA
AC P42047;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PHOTOSYSTEM I REACTION CENTRE SUBUNIT IV (PHOTOSYSTEM I 19.5 KDA
DE PROTEIN) (PSI-E) (PS I SUBUNIT 6) (FRAGMENT).
GN PSAE.
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;

```

OC Cucurbitales; Cucurbitaceae; Cucumis.
RN [1]
RP SEQUENCE.
RC TISSUE=COLEYEDON.
RX MEDLINE-91355209; PubMed-1883835;
RA Iwasaki Y., Ishikawa H., Hibino T., Takabe T.;
RT "Characterization of genes that encode subunits of cucumber PS I
RL complex by N-terminal sequencing."
RL Biochim. Biophys. Acta 1059:141-148(1991).
CC -1- FUNCTION: MAY FORM COMPLEXES WITH FERREDOXIN AND FERREDOXIN-
CC OXIDOREDUCTASE IN PHOTOSYSTEM I (PS I) REACTION CENTRE.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE PSAE FAMILY.
KM Photosynthesis; Photosystem I; Chloroplast; Thylakoid membrane.
FT NON_TER 19
SQ SEQUENCE 19 AA; 1810 MW; A93E8BD089FB738 CRC64;

Query Match 19.4%; Score 20; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAT 12
DB 8 PAAT 11

RESULT 23
COXM_THUOB STANDARD; PRT; 20 AA.
ID COXM_THUOB
AC P80981;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VIIB-LIVER (EC 1.9.3.1) (FRAGMENT).
OS Thunnus obesus (Bigeye tuna).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
CC Scombridae; Thunnus.
RN [1]
RP SEQUENCE.
RC TISSUE=LIVER.
RX MEDLINE-97454291; PubMed-9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver."
RL Eur. J. Biochem. 248:99-103(1997).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
CC 4 FERRICYTOCHROME C.
KM Oxidoreductase; Mitochondrion.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2213 MW; 03B902A70F3E41A6 CRC64;

Query Match 19.4%; Score 20; DB 1; Length 20;
Best Local Similarity 75.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 12 THOD 15
DB 4 SHOD 7

RESULT 24
COXM_THUOB STANDARD; PRT; 20 AA.
ID COXM_THUOB
AC P80980;
DT 01-NOV-1997 (Rel. 35, Created)

```

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VIIB-HEART (EC 1.9.3.1) (FRAGMENT).
OS Thunnus obesus (Bigeye tuna).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
CC Scombridae; Thunnus.
RN [1]
RP SEQUENCE.
RC TISSUE=HEART.
RX MEDLINE-97454291; PubMed-9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver."
RL Eur. J. Biochem. 248:99-103(1997).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
CC 4 FERRICYTOCHROME C.
KM Oxidoreductase; Mitochondrion.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2303 MW; 0A33BD340065A6 CRC64;

Query Match 19.4%; Score 20; DB 1; Length 20;
Best Local Similarity 75.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 12 THOD 15
DB 4 SHOD 7

RESULT 25
PGK_CLOPA STANDARD; PRT; 20 AA.
ID PGK_CLOPA
AC P81346;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE POTASSIUM PHOSPHOGLYCERATE KINASE (EC 2.7.2.3) (CP 11) (FRAGMENT).
GN PGK.
OS Clostridium pasteurianum.
CC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
CC Clostridium.
RN [1]
RP SEQUENCE.
RC STRAIN=W5;
RX MEDLINE-98291870; PubMed-9629918;
RA Flensburg R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5."
RL Electrophoresis 19:802-806(1998).
CC -1- CATALYTIC ACTIVITY: ATP + 3-PROSPHO-D-GLYCERATE -> ADP +
CC 3-PROSPHO-D-GLYCEROL PHOSPHATE.
CC -1- PATHWAY: SECOND STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC 5.6, ITS MW IS: 56.2 KDA.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
DR INTERPRO: IPR001576;
DR PROSITE: PS00111; PGLYCERATE KINASE, PARTIAL.
KM Transferase; Kinase; Glycolysis.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2355 MW; 749E31E595C85529 CRC64;

Query Match 19.4%; Score 20; DB 1; Length 20;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;

```

Tue Feb 6 08:48:26 2001

us-08-981-824-7.rsp

Page 9

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 THODID 17
|:|:|
Db 6 TIEDID 11

Search completed: February 5, 2001, 10:56:04
job time: 504 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:49:32 ; Search time 58.67 seconds
(without alignments)
23.147 Million cell updates/sec

Title: US-08-981-824-7

Sequence: 103
1 FFRWVSNPAATHQDIDFLI 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 segs, 67900655 residues

Total number of hits satisfying chosen parameters: 5001

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	31	30.1	16	2	S02473	coat protein VP1 -
2	30	29.1	20	2	S00494	hemocyanin chain I
3	27	26.2	22	2	S35200	hypothetical prote
4	25	24.3	14	2	S62374	alpha-1-antichymo
5	25	24.3	15	2	A45096	thyrotropin-releas
6	25	24.3	18	2	S70611	30K protein - Enge
7	25	24.3	21	2	A59335	probable bacteriop
8	25	24.3	24	2	S62890	non-histone protei
9	25	24.3	25	2	B57001	endo-1,4-beta-xyla
10	24	23.3	8	2	I49404	prealbumin - weste
11	24	23.3	15	2	PQ0194	S2-glycoprotein
12	24	23.3	17	2	PH1312	Ig heavy chain DJ
13	24	23.3	18	2	A54195	Na+/K+-exchanging
14	24	23.3	19	2	S00495	hemocyanin chain I
15	24	23.3	23	2	PH1364	Ig heavy chain DJ
16	24	23.3	24	1	S58242	pyrroloquinoline q
17	24	23.3	24	1	I54329	gene NF1 protein -
18	24	23.3	25	2	JP0045	ribosomal protein
19	23	22.3	15	2	PQ0174	styliar glycoprotei
20	23	22.3	22	2	PH1335	Ig heavy chain DJ
21	23	22.3	22	2	H49410	t-complex polypept
22	23	22.3	23	2	B43958	GTP-binding protei
23	23	22.3	24	2	PT0076	leucine:dehydrogen
24	23	22.3	25	2	T07504	hypothetical prote
25	23	22.3	25	2	S35926	T-cell receptor ga
26	22	21.4	11	2	S42449	anti protein - pha
27	22	21.4	12	2	S43013	hypothetical prote
28	22	21.4	12	2	J02308	hypothetical 1.4K
29	22	21.4	12	2	J02318	hypothetical 1.4K

30	22	21.4	16	2	PH0763	T-cell receptor be
31	22	21.4	17	2	S58660	H+-transporting AT
32	22	21.4	17	2	B48943	phage antigenic de
33	22	21.4	18	2	I52851	brain-derived neur
34	22	21.4	19	2	PA0010	seed storage prote
35	22	21.4	19	2	S03519	T-cell receptor ga
36	22	21.4	20	2	S72501	protein kinase C 1
37	22	21.4	20	2	S59494	formate dehydrogen
38	22	21.4	21	2	B35417	30K serine protei
39	22	21.4	21	2	A47144	aquaporin-CHIP - h
40	22	21.4	21	2	S27039	blibrotaxin - Bbro
41	22	21.4	23	2	A60226	pyruvate dehydroge
42	22	21.4	24	2	I39289	ZF3 domain - human
43	22	21.4	24	2	PN0456	basic eosinophil p
44	21.5	20.9	20	2	B60365	chymotrypsin inh
45	21	20.4	13	2	A54326	glandular kallikre

ALIGNMENTS

RESULT 1
S02473
coat protein VP1 - human poliovirus 1 (fragment)
C:Species: human poliovirus 1
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 18-Jun-1993
C:Accession: S02473
R:Martin, A.; Wychowski, C.; Couderc, T.; Craignic, R.; Hogle, J.; Girard, M.
EMBO J. 7, 2839-2847, 1988
A:Title: Engineering a poliovirus type 2 antigenic site on a type 1 capsid results in
A:Reference number: S02473; PMID:89030650
A:Accession: S02473
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-16 <MAR>

Query Match 30.1% Score 31; DB 2; Length 16;
Best Local Similarity 55.6% Pred. No. 40;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 MYISNPAAT 12
DB 1 MYVDNPAST 9

RESULT 2
S00494
hemocyanin chain II - Japanese spiny lobster (fragment)
C:Species: Panulirus japonicus (Japanese spiny lobster)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 01-Aug-1997
C:Accession: S00494
R:Makino, N.; Kimura, S.
Eur. J. Biochem. 173, 423-430, 1988
A:Title: Subunits of Panulirus japonicus hemocyanin. 1. Isolation and properties.
A:Reference number: S00492; PMID:88196131
A:Accession: S00494
A:Molecule type: protein
A:Residues: 1-20 <MAK>
C:Superfamily: hemocyanin
C:Keywords: copper; hemolymph; hexamer; oxygen carrier

Query Match 29.1% Score 30; DB 2; Length 20;
Best Local Similarity 41.2% Pred. No. 78;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 MYISNPAATHQDIDFLI 20
DB 2 VVASSTAHKQODINHL 18

RESULT 3

S35200
 hypothetical protein 7 - Saccharopolyspora hirsuta
 C:Species: Saccharopolyspora hirsuta
 C>Date: 10-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 17-Mar-1999
 C:Accession: S35200
 R:Le Gouill, C.; Desmarais, D.; Dery, C.V.
 M61. Gen. Genet. 240, 146-150, 1993
 A:Title: Saccharopolyspora hirsuta 367 encodes clustered genes similar to ketoacyl synth
 A:Reference number: S35194; MUID:93344153
 C:Accession: S35200
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-22 <GOU>

Query Match 26.2%; Score 27; DB 2; Length 22;
 Best Local Similarity 62.5%; Pred. No. 2.9e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 RMVSNPA 10
 DB 12 RLVVSAPA 19

RESULT 4
 S62374
 alpha-1-antichymotrypsin - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 31-Mar-1997
 C:Accession: S62374
 R:Tsuda, M.; Sei, T.; Ohkubo, T.; Yamamura, M.; Kamiguchi, H.; Akatsuka, A.; Tsuda, T.;
 Eur. J. Biochem. 235, 821-827, 1996
 A:Title: The defective secretion of a naturally occurring alpha-1-antichymotrypsin varia
 A:Reference number: S62374; MUID:96184564
 C:Accession: S62374
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-14 <TSU>

Query Match 24.3%; Score 25; DB 2; Length 14;
 Best Local Similarity 44.4%; Pred. No. 3.7e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 FFRWISNP 9
 DB 3 FFMKVTNP 11

RESULT 5
 A45096
 thyrotropin-releasing hormone receptor, splice form 387 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
 C:Accession: A45096
 R:de la Pena, P.; Delgado, L.M.; del Camino, D.; Barros, F.
 J. Biol. Chem. 267, 25703-25708, 1992
 A:Title: Two isoforms of the thyrotropin-releasing hormone receptor generated by alterna
 A:Reference number: A45096; MUID:93100278
 C:Accession: A45096
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-15 <DEI>
 A:Cross-references: GB:S51512; NID:9261982; PIDN:AAB24549.1; PID:9261983
 A:Experimental source: GH3 anterior pituitary cells
 A:Note: sequence extracted from NCBI backbone (NCBIP:120927)

Query Match 24.3%; Score 25; DB 2; Length 15;
 Best Local Similarity 44.4%; Pred. No. 4e+02;
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 8 NPATRHDI 16

DB 6 DPLVHMDI 14

RESULT 6
 S70611
 30K protein - Engelmannia pinnatifida (fragment)
 C:Species: Engelmannia pinnatifida
 C>Date: 14-Apr-1998 #sequence_revision 08-May-1998 #text_change 07-May-1999
 C:Accession: S70611
 R:Huynh, Q.K.; Borgmeyer, J.R.; Smith, C.E.; Bell, L.D.; Shah, D.M.
 Biochem. J. 316, 723-727, 1996
 A:Title: Isolation and characterization of a 30 kDa protein with antifungal activity
 A:Reference number: S70611; MUID:96265034
 C:Accession: S70611
 A:Molecule type: protein
 A:Residues: 1-18 <HUY>
 A:Experimental source: leaves
 C:Keywords: antifungal

Query Match 24.3%; Score 25; DB 2; Length 18;
 Best Local Similarity 40.0%; Pred. No. 5e+02;
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 FFRWISNP 10
 DB 7 FFLALQXPA 16

RESULT 7
 A59325
 Probable bacteriophage receptor Bacta [Imported] - Bacteroides fragilis (fragment)
 C:Species: Bacteroides fragilis
 C>Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000
 C:Accession: A59325
 R:Frias-Lopez, J.
 submitted to the Protein Sequence Database, July 2000
 A:Description: Identification of cell wall proteins of Bacteroides fragilis to which
 A:Reference number: A59325
 C:Accession: A59325
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-21 <FRI>
 A:Experimental source: strain ATCC 51477
 A:Note: putative receptor for bacteriophage B40-8

Query Match 24.3%; Score 25; DB 2; Length 21;
 Best Local Similarity 50.0%; Pred. No. 6e+02;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 9 PAATHQDI 16
 DB 9 PAALHDEM 16

RESULT 8
 S62890
 non-histone protein 1 75K chain - human (fragments)
 C:Species: Homo sapiens (man)
 C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
 C:Accession: S62890
 R:Oderwald, H.; Hughes, M.J.; Jost, J.P.
 FEBS Lett. 382, 313-318, 1996
 A:Title: Non-histone protein 1 (NHP1) is a member of the Ku protein family which is u
 A:Reference number: S62889; MUID:96184511
 C:Accession: S62890
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-14;15-24 <ODE>

Query Match	24.3%	Score 25	DB 2	Length 24
Best Local Similarity	33.3%	Pred. No. 7e+02		
Matches	4	Conservative	3	Mismatches 5
				Indels 0
				Gaps 0

RESULT 9
B57001
endo-1,4-beta-xylosylase (EC 3.2.1.8) 2 - Streptomyces roseiscleroticus (fragment)
C:Species: Streptomyces roseiscleroticus
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 06-Dec-1996
C:Accession: B57001
R:GrubbsRL, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
Protein Expr. Purif. 4, 120-129, 1993
A:Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxylanases
A:Reference number: A57001; MUID:93229899
A:Accession: B57001
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <GRA>
A:Experimental source: strain NRRL B-11019
A:Note: sequence extracted from NCBI backbone (NCBIP:130008)
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: endo-1,4-beta-xylosylase; endo-1,4-beta-xylosylase homology
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation

Query Match	24.3%	Score 25	DB 2	Length 25
Best Local Similarity	36.4%	Pred. No. 7.4e+02		
Matches	4	Conservative	3	Mismatches 4
				Indels 0
				Gaps 0

RESULT 10
149404
prealbumin - western wild mouse (fragment)
C:Species: Mus spretus (western wild mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I49404
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Mezaki, Y.; Nadeau, J.H.
Mmm. Genome 5, 349-355, 1994
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A:Reference number: I48934; MUID:94319082
A:Accession: I49404
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-8 <RES>
A:Cross-references: EMBL:U05668; NID:g497008; PIDN:AAB60461.1; PID:g642825

Query Match	23.3%	Score 24	DB 2	Length 8
Best Local Similarity	80.0%	Pred. No	2e+05	
Matches	4	Conservative	1	Mismatches 0
				Indels 0
				Gaps 0

RESULT 11
P00194
52-glycoprotein - Persian tobacco (fragment)
C:Species: Nicotiana glauca (Persian tobacco)
C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 03-Mar-1995
C:Accession: P00194

R:Manu, S.L.; Williams, E.G.; Atkinson, A.; Anderson, M.A.; Cornish, E.C.; Grego, B.;
Planta 169, 184-191, 1986
A:Title: Style proteins of a wild tomato (*Lycopersicon peruvianum*) associated with ex-
A:Reference number: P00192
A:Accession: P00194
A:Molecule type: protein
A:Residues: 1-15 <MAU>
A:Experimental source: style
C:Keywords: glycoprotein

	Query Match	23.3%	Score 24:	DB 2:	Length 15:
	Best Local Similarity	27.3%	Pred.	6e+02:	
Matches	3; Conservative	6;	Mismatches	2;	Indels
				0;	Gaps 0;
QY	1 FFRWYSNPAA	11	:	:::::	:
DbB	4 YKQLVLTXPAS	14			

RESULT 12
PH1312
1g heavy chain DJ region (clone C68-105) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1312
R:Masserman, R.; Gallili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type D_H joining in young children with B precursor lymphocytes
A:Reference number: PH1302; MUID:93094761
A:Accession: PH1312
A:Molecule type: DNA
A:Residues: 1*17 <MAS>
C:Keywords: heterodimer; immunoglobulin

Query Match	23.3%	Score 24:	DB 2:	length 17:
Best Local Similarity	54.5%	Pred No	6.9e+02:	
Matches	6;	Conservative	1;	Mismatches 4; Indels 0; Gaps 0

```

RESULT 13
A54195
Na+/K+-exchanging ATPase (EC 3.6.1.37) - spiny dogfish (fragment)
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 13-oct-1994 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C:Accession: A54195
R:Esmann, M.; Aralish, S.J.; Solttrup-Jensen, L.; Marsh, D.
Biochemistry 33, 8044-8050, 1994
A:Title: Structural integrity of the membrane domains in extensively tyrosylated Na,K
A:Reference number: A54195; MUID:94297020
A:Accession: A54195
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-18 <ESM>
A:Experimental source: rectal gland
A:Note: sequence extracted from NCBI backbone (NCBIP:149363)
C:Keywords: hydrolase

```

Query Match	23.38%	Score 24;	DB 2;	length 18;
Best Local Similarity	38.58%	Pred. No. 7.4e+02;		
Matches	5;	Conservative	2;	Mismatches 6; Indels 0; Gaps 0
QY	6	ISNPATHQDIDF	18	
	::			
Db	6	LTPPTPTPYDIKF	18	

RESULT 14

S00495

hemocyanin chain III - Japanese spiny lobster (fragment)

C:Species: Panulirus japonicus (Japanese spiny lobster)

C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 01-Aug-1997

C:Accession: S00495

R: Makino, N.; Kimura, S.

Eur. J. Biochem. 173, 423-430, 1988

A:Title: Subunits of Panulirus japonicus hemocyanin. 1. Isolation and properties.

A:Reference number: S00492; MUID:88196131

A:Accession: S00495

A:Molecule type: protein

A:Residues: 1-19 <MAK>

C:Superfamily: hemocyanin

C:Keywords: copper; hemolymph; hexamer; oxygen carrier

Query Match 23.3%; Score 24; DB 2; Length 19;

Best Local Similarity 42.9%; Pred. No. 7.9e+02;

Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 SNPAATHODIDFLI 20

DB 5 SCNAHQDIDNHIL 18

RESULT 15

PH1364

lg heavy chain DJ region (clone C178-136A) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1364

R: Masserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph

A:Reference number: PH1302; MUID:93094761

A:Accession: PH1364

A:Molecule type: DNA

A:Residues: 1-23 <MAS>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 23.3%; Score 24; DB 2; Length 23;

Best Local Similarity 55.6%; Pred. No. 9.9e+02;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 SNPAATHOD 15

DB 1 STPAATHNRD 9

RESULT 16

S58242

pyrroloquinoline quinone precursor pqqa - Pseudomonas fluorescens

N:Alternate names: pyrroloquinoline quinone biosynthesis A

C:Species: Pseudomonas fluorescens

C>Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000

C:Accession: S58242

R:Schneider, U.; Keel, C.; Defago, G.; Haas, D.

submitted to the EMBL Data Library, May 1995

A:Description: In5-directed cloning of pqg genes from Pseudomonas fluorescens CHAO: the1

A:Reference number: S58239

A:Accession: S58242

A:Molecule type: DNA

A:Residues: 'MYROHSHPOBSNF', 1-24 <SCH>

A:Cross-references: EMBL:X87299; NID:9929799; PIDN:CA60731.1; PID:9929803; PIDN:CA6073

A>Note: In Genbank entry PF000ABC, release 116.0, the indicated alternative sequence (F

C:Genetics:

A:Gene: pqqa

C:Superfamily: pyrroloquinoline quinone precursor pqqa

C:Keywords: quinoprotein

F:16/20/Product: pyrroloquinoline quinone #status predicted <MAT>

F:16-20/Cross-link: pyrroloquinoline quinone (Glu, Tyr) #status predicted

Query Match 23.3%; Score 24; DB 1; Length 24;

Best Local Similarity 46.7%; Pred. No. 1e+03;

Matches 7; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 MWISNPATHDIDFL 18

DB 1 MTWSKPAYTDLRIGF 15

RESULT 17

I54329

gene NF1 protein - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 05-Nov-1999

C:Accession: I54329

R:Upadhyaya, K.; Shen, M.; Cherryson, A.; Farnham, J.; Maynard, J.; Huson, S.M.; Harp

Hum. Mol. Genet. 1, 735-740, 1992

A:Title: Analysis of mutations at the neurofibromatosis 1 (NF1) locus.

A:Reference number: I54329; MUID:93258316

A:Accession: I54329

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-24 <RES>

A:Cross-references: GB:S61428; NID:9300449; PIDN:AAB26693.1; PID:9300450

C:Genetics:

A:Gene: GDB:NF1

A:Cross-references: GDB:I20231; OMIM:162200

A:Map position: 17q11.2-17q11.2

Query Match 23.3%; Score 24; DB 2; Length 24;

Best Local Similarity 100.0%; Pred. No. 1e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 THQD 15

DB 6 THQD 9

RESULT 18

JP0045

ribosomal protein L30 - Staphylococcus aureus (fragment)

C:Species: Staphylococcus aureus

C>Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 03-May-1996

C:Accession: JP0045

R:Ochi, K.

submitted to JRPID, February 1994

A:Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal

A:Reference number: JP0042

A:Accession: JP0045

A:Molecule type: protein

A:Residues: 1-25 <OCH>

C:Superfamily: Escherichia coli ribosomal protein L30

C:Keywords: protein biosynthesis; ribosome

Query Match 23.3%; Score 24; DB 2; Length 25;

Best Local Similarity 29.4%; Pred. No. 1.1e+03;

Matches 5; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 RWVISNPATHDIDFL 19

DB 9 RSVIGRPETQRKTVEAL 25

RESULT 19

PQ0174

stylar glycoprotein 3 - Persian tobacco (fragment)

N:Alternate names: glycoprotein S7

C:Species: Nicotiana glauca (Persian tobacco)

C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 13-Mar-1997
 C:Accession: P00174
 R:Jahnen, W.; Batterham, M.P.; Clarke, A.E.; Moritz, R.L.; Sampson, R.
 Plant Cell 1, 493-499, 1989
 A:Title: Identification, isolation, and N-terminal sequencing of style glycoproteins ass
 A:Reference number: P00173; MUID:92404717
 A:Accession: P00174
 A:Molecule type: protein
 A:Residues: 1-15 <JAH>
 A:Experimental source: style
 C:Comment: The protein is involved in self-incompatibility of flowering plants.
 C:Superfamily: Enterobacter ribonuclease
 C:Keywords: glycoprotein

Query Match 22.3% Score 23; DB 2; Length 15;
 Best Local Similarity 36.4% Pred. No. 8.9e+02;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 FFRVVISNPA 11
 : : : : :
 Db 4 YMOGLVQWPA 14

RESULT 20
 PH1325
 Ig heavy chain DJ region (clone C199-121) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C:Accession: PH1325
 R:Wasserman, R.; Gailili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
 J. Exp. Med. 176, 1577-1581, 1992
 A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph
 A:Reference number: PH1302; MUID:93094761
 A:Accession: PH1325
 A:Molecule type: DNA
 A:Residues: 1-22 <WAS>
 C:Superfamily: Immunoglobulin V region: Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin

Query Match 22.3% Score 23; DB 2; Length 22;
 Best Local Similarity 44.4% Pred. No. 1.4e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 VISNPATN 13
 : : : : :
 Db 2 IVVVPAAIH 10

RESULT 21
 H49410
 t-complex polypeptide 1 homolog (peak 7 fraction) - rabbit (fragments)
 N:Alternate names: chaparrin homolog (peak 7)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 23-Mar-1995
 C:Accession: H49410
 R:Rommlaere, H.; Van Troys, M.; Gao, Y.; Melki, R.; Cowan, N.J.; Vandekerckhove, J.; An
 Proc. Natl. Acad. Sci. U.S.A. 90, 11975-11979, 1993
 A:Title: Eukaryotic cytosolic chaparrin contains t-complex polypeptide 1 and seven relat
 A:Reference number: A49410; MUID:94085752
 A:Accession: H49410
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-22 <ROM>
 A:Experimental source: reticulocyte

Query Match 22.3% Score 23; DB 2; Length 22;
 Best Local Similarity 44.4% Pred. No. 1.4e+03;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 WVISNPAT 12

Db 1 LEVTNDPAT 9
 : : : : :
 : : : : :

RESULT 22
 B43958
 GTP-binding protein, synaptic vesicle-specific - marbled electric ray (fragments)
 C:Species: Torpedo marmorata (marbled electric ray)
 C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 28-May-1999
 C:Accession: B43958
 R:Volkmann, W.; Pensner, J.; Elferink, L.A.; Schilling, J.; Scheller, R.H.
 Brain Res. Mol. Brain Res. 11, 285-290, 1991
 A:Title: A synaptic vesicle specific GTP-binding protein from ray electric organ.
 A:Reference number: A43958; MUID:92099847
 A:Accession: B43958
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-23 <VOL>
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology
 C:Keywords: GTP binding

Query Match 22.3% Score 23; DB 2; Length 23;
 Best Local Similarity 30.0% Pred. No. 1.5e+03;
 Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 11 AFHODIDFLI 20
 : : : : :
 Db 2 AVDQNFDMV 11

RESULT 23
 PT0076
 leucine dehydrogenase (EC 1.4.1.9) - Corynebacterium diptheriae (fragment)
 C:Species: Corynebacterium diptheriae
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
 C:Accession: PT0076
 R:Misono, H.; Sugihara, K.; Kuwamoto, Y.; Nagata, S.; Nagasaki, S.
 Agric. Biol. Chem. 54, 1491-1498, 1990
 A:Title: Leucine dehydrogenase from Corynebacterium pseudodiptheriticum: Purificatio
 A:Reference number: PT0076; MUID:91103959
 A:Accession: PT0076
 A:Molecule type: protein
 A:Residues: 1-24 <MIS>
 A:Experimental source: strain ICR2210
 C:Comment: This enzyme catalyzes the reversible deamination of branched-chain and str
 C:Keywords: oxidoreductase

Query Match 22.3% Score 23; DB 2; Length 24;
 Best Local Similarity 66.7% Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 12 THDID 17
 : : : : :
 Db 4 THVDID 9

RESULT 24
 T07504
 hypothetical protein 25 - Japanese black pine chloroplast
 C:Species: chloroplast Pinus thunbergiana (Japanese black pine)
 C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 18-Aug-2000
 C:Accession: T07504
 R:Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Suglura, M.
 Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
 A:Title: Loss of all ndh genes as determined by sequencing the entire chloroplast gen
 A:Reference number: Z16030; MUID:95024047
 A:Accession: T07504
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-25 <WAK>
 A:Cross-references: EMBL:DJ7510; MID:9529643; PIDN:BA04382.1; PID:91262665

C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 22.3%; Score 23; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 4 MVTSP 9
Db 15 MVTSP 20

RESULT 25

S35926
T-cell receptor gamma chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 30-May-1997
C:Accession: S35926
R:Mathioudakis, G.; Platsoucas, C.D.
submitted to the EMBL Data Library, May 1993
A:Description: An alternative splicing between Vgamma2gamma2.3 and Cgamma2 gene segments
A:Reference number: S32764
A:Accession: S35926
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-25 <MAT>
A:Cross-references: EMBL:Z22688
C:Keywords: T-cell receptor

Query Match 22.3%; Score 23; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 FMVVISNPATHODI 16
Db 7 YLLIRNSLAEQHL 21

Search completed: February 5, 2001, 10:49:32
Job time: 745 sec